

75938

STIC-Biotech/ChemLib

**From:** Fredman, Jeffrey  
**Sent:** Tuesday, September 17, 2002 2:36 PM  
**To:** STIC-Biotech/ChemLib  
**Cc:** Leffers, Gerald  
**Subject:** FW: 09/846,456

PLEASE RUSH.

I Approve.

Jeff Fredman

-----Original Message-----

**From:** Leffers, Gerald  
**Sent:** Tuesday, September 17, 2002 2:27 PM  
**To:** Fredman, Jeffrey  
**Subject:** RE: 09/846,456

Hi Jeff, I'm not restricting because 1) the other sequences are all (supposedly) comprised within the whole ABC1 gene of SEQ ID NO: 1 (i.e. SEQ ID NO: 1 is a genomic gene sequence; SEQ ID NO: 3 corresponds to the first exon, SEQ ID NO: 5 to the second exon, etc.) and 2) the case has already been restricted by another examiner without restriction based upon these sequences (the case is at the stage of election by applicants). Thanks, Gerry

*Gerald G. Leffers Jr., PhD*  
 Examiner, Art Unit 1636  
 Crystal Mall 1, Room 11A09  
 703-308-6232

-----Original Message-----

**From:** Fredman, Jeffrey  
**Sent:** Tuesday, September 17, 2002 2:24 PM  
**To:** Leffers, Gerald  
**Subject:** RE: 09/846,456

Gerald,

Given this situation, why are you not restricting between the sequences and requiring election of a single DNA sequence?

Jeff

-----Original Message-----

**From:** Leffers, Gerald  
**Sent:** Tuesday, September 17, 2002 2:05 PM  
**To:** Fredman, Jeffrey  
**Subject:** 09/846,456

Hi Jeff, please approve a RUSH search for SEQ ID NOS: 1-5 of this application. Total DNA is ~7 kb. Claims are directed to the whole sequence, or alternatively, a polynucleotide comprising 20 consecutive nucleotides of any one of SEQ ID NOS: 1-5. SEQ ID NOS: 3-5 should be comprised within SEQ ID NO: 1. I have read through the specification and have not been able to find if SEQ ID NO: 2 is found within SEQ ID NO: 1. It may be that someone could do a quick alignment of sequences in the case and determine if a single oligo search of SEQ ID NO: 1 would cover each of the other claimed sequences. Thanks, Gerry

Searcher: <u>D. Schreiber</u>	TYPE OF SEARCH:	VENDOR/COST (where applic.)
Phone: <u>708-4292</u>	NA Sequences: <u>10</u>	STN: _____
Location: <u>CM 6A03</u>	AA Sequences: _____	DIALOG: _____
Date Picked Up: <u>9/18</u>	Structures: _____	Questel/Orbit: _____
Date Completed: <u>9/20</u>	Bibliographic: _____	DRLink: _____
Searcher Prep/Review: <u>14</u>	Litigation: _____	Lexis/Nexis: _____
Clerical: _____	Full text: _____	Sequence Sys.: <u>CompuGen</u>
Online time: <u>8</u>	Patent Family: _____	WWW/Internet: _____
	Other: _____	Other (specify): _____

24 30

12



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 19, 2002, 23:21:26 ; Search time 5225.75 Seconds  
(without alignments)  
12938.552 Million cell updates/sec

Title: US-09-846-456-1

Perfect score: 3231

Sequence: 1 acaggcgtggtggcaggtg.....gccccacatccccaccactt 3231

Scoring table:

IDENTITY\_NUC  
Gap 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_hgt.\*

3: gb\_in.\*

4: gb\_em.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sy.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sy.\*

28: em\_un.\*

29: em\_vi.\*

30: em\_hgt\_hum.\*

31: em\_hgt\_inv.\*

32: em\_hgt\_other.\*

33: em\_hgtg\_inv.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
------------	-------	-------	--------	----	----	-------------

#### ALIGNMENTS

RESULT 1

AX351029

LOCUS

DEFINITION

AX351029

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

BASE COUNT

ORIGIN

AX351029

Sequence

AX351029

AX351029.1

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Homidae; Homo.

1 (sites)

Rosier-Montus, M.F., Prades, C., Lemoine, C., Naudin, L., Deneffe, P.,

Brewer, B., Duverger, N., Remaley, A. and Santamarina-Fojo, S.

Regulatory nucleic acid sequences of the abcl gene

Patent: WO 0183746-A 1 08-NOV-2001;

Aventis Pharma S.A. (FR)

Location/Qualifiers

1..3231

/organism="Homo sapiens"

/db\_xref="taxon:9606"

809 a

773 c

876 g

773 t

DNA

3231 bp

1 from Patent WO0183746.

GI:18616385

linear

PAT 06-FEB-2002

AX351029

Sequence

AX351029

AX351029

AX351029

AX351029

AX351029

AX351029

AX351029

AX351029

AX351029

AX351029

AX351029

AX351029

AX351029

AX351029

AX351029

AX351029

AX351029

AX351029

AX351029

Query Match	100.0%;	Score 3231;	DB 6;	Length 3231;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3231;	Conservative	0;	Mismatches	0;
Indels	0;	Gaps		
QY	1	acaggcctggtgagcaggtgctgtaactcagttactcggaggtcgaggtgcaatga	60	
DB	1	ACAGGCGATGGTGGCAGTGCCCTGTANYCTCAGTTACTCGGGAGGTGGAGGTGGCAATGA	60	
QY	61	gcccagatgcgcaccattgcactccagctcgggcaacaagaagtgaactccatctcaatt	120	
DB	61	GCCCAGATGCGCACCATATGCACATCCAGCTGGGCCAACAAAAGGTGAACATCCATCTCAATT	120	
QY	121	aaaaaaaaagaatatttgctgctcagcttcaaataggtaggagaagaagagagagg	180	
DB	121	AAAAAAAAGAAGTAATTTTGGTGTGCACCTCAATAGGTAGGAGAAGAGGAGAGG	180	
QY	181	agatggaggttcagggagatctaattactctcaaaatcatgtaggaagataaacact	240	
DB	181	AGATGGAGGCTCAGGAGATCAATTAAGTCTCTAAATCATGCTAGGAAGATAAACCT	240	
QY	241	ttaataaacactctctgtttataacatcatctgcgaaggagctcaaaagtttcaaca	300	
DB	241	TTTAATAACACATCTCTGCTTTTATAACATCATCTGCCAAGGAGCTCAAGAGTTTCAACA	300	
QY	301	aagttcacctttcagaaaaaccccttgaggaagacagaataacatctctctccatttta	360	
DB	301	AAGTTCACCTTTCAGAAAAACCCCTTTGAGGAAGACAGATAATACATCTTCTCTCCATTTTA	360	
QY	361	aagatgaagaacacgcccgggcacaaatggcttaatccctgtaatcccagcactttggagg	420	
DB	361	AAGATGAAGAACAACGCCGGGACAAATGGCTAATGCCCTGTATCCACGACCTTTGGGAGG	420	
QY	421	ctgagccagagatcgccttgagctccagagtttgagaccagcctggataacatggcaaa	480	
DB	421	CTGAGGCCAGAGATCGCTTGAGCTCCAGAGTTTCAGACCCAGCCTGGATATACATGGCAAA	480	
QY	481	acctgtctctcaaaaaaaatcacaaaattagatgggtgtggtggcgatgcacctgtggt	540	
DB	481	ACCTGTCTCTCAAAAAAAATACAAAAATAGATGGGTGTGGTGGCATGCACTGTGGT	540	
QY	541	ccagctacttggaggctaaagtggaggatcgctctagcccaggagatcaagtctaca	600	
DB	541	CCCAGCTACTTGGGAGGCTAAGGTGGGAGATCGCTCTGAGCCAGGGAGTCAAGTCTACA	600	
QY	601	ctgagccatgattgatcactcagctccagcctgggttagacagagcaagacctgtctca	660	
DB	601	CTGAGCCATGATTGGATCCTCACTCCAGCCTGGGTAGACAGAGCAAGACCTGTCTCA	660	
QY	661	aaaaaagaatgaaagagaaagaagaagagagagagagagatgagggagagg	720	
DB	661	AAAAAAGAAATGAAGAAGAAAGAAAGAGAGAGAGAGAGAGATGAGGGGAGGAGG	720	
QY	721	gaggggggaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaagga	780	
DB	721	GAGGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	780	
QY	781	gatgaacagagcgagaagacttacgtaaattgctcatctggttggttgcaagtttga	840	
DB	781	GATGAACAGAGCGCAGAAAGACTTTACGTAAATTTGCTCATCTGTTGTGTCAAGTTGA	840	
QY	841	ccccaaaaccatattattgaccaaggttattctttgactgagcgaaggggtccgctct	900	
DB	841	CCCCAAAACCCAAATTTATGACCAAGGTTATTCTTTTGACTGAGCGCAGGGGTCCGCTCT	900	
QY	901	cctgggccttggcctttagaagctcatctctggccttctgagatccatcccttcttt	960	
DB	901	CCTGGCCTTGGCCTTAGAAAGCTCATCTCTGGCCTTCTGAGATCCATCCCTTCTTT	960	
QY	961	ttattttttgacacgagcttgtctctgtctgtoactcaggtcgagtgagtgcatgatc	1020	
DB	961	TTATTTTCTTGACACGAGCTTCTGCTGTGCATCTCAGCTGGAGTGCAGTGGCATGATC	1020	

Qy	1021	tcgactcactgtaacctctgctcccggttccaagcgattctcctgcctcaagcctcctga	1081
Db	1021	TCGACTCACTGTAACTCTGCGCTCCCGGTTCAAGCGAATTCTCTGCTCAGCCTCTCTGA	1081
Qy	1081	gataacaggccgcgcacacacatctgactaattttatttattttagtaaagactgggtt	1140
Db	1081	GATAACAGGCGCGCCGCCACCATCTGGCTAATTTTGTATTTTGTAAAGACTGGGTT	1140
Qy	1141	tcatcatgttgccaggttggtttcgaactcctgacctgaggtgagctgccaccttggc	1200
Db	1141	TCATCATGTGCGCAGGTGGTTGTTGAACTCTGCACCTGAGGTGAGCTGCCACCTTGGC	1200
Qy	1201	ctccaaaagtctgggattacaagcatgagccaactgcccagctcagatccatcccttt	1260
Db	1201	CTCCAAAAGTCTGGGATTACAGGATGAGCCACATGCGCCAGCTCAGATCCATCCCTTT	1260
Qy	1261	ctaaggggcaaacagtcctcattggttgcaaaaggggcattgccaccagagttatgagtacctgg	1320
Db	1261	CTAAGGGCAACAGTCCATGGTGCAAGGGGCCATGCGCACACAGAGTTATGAGTACCTGG	1320
Qy	1321	gactccagaattccttgctgctggctccacatgcacttccagggcctgcttggggctc	1380
Db	1321	GACTCCAGAAATCCTTGTGCTGTGGCGCTCCATGCATCTCCAGGGCGCTGCTGGGGCTC	1380
Qy	1381	tctatgctctgctcctgagtttgatgaaccactgatgtgagtcactggccttgagcc	1440
Db	1381	TTCTATGCGTCTGCTCGAGTGTGATAGAACCACTGATGTGAGTACCTGGGCTTGAGCC	1440
Qy	1441	gtggcctggagatcctctgtgactgtagcatggaggggctgtgcagctgaatgctgcga	1500
Db	1441	GTGGCCTGGAGATCCCTGTGTGACTGTAGCATGGAGGGGCTGTGCAGCTGAATGCTGCA	1500
Qy	1501	tgcagtggtgggagttctggaatgatggagctggaggtggaagagaaagttagccttg	1560
Db	1501	TGCAGGTGGTGGGAGTTCGGAATATGATGGAGCTGGAGCTGGGAAGAGATGAGGCTTG	1560
Qy	1561	gggcagctctctcatgccacctcattcttggccaaactcaggtccaaactgtgaagagctc	1620
Db	1561	GGGCAGCTCTCTCATGCCACCTCATTTCTGGCCAAAACCTCAGGTCAAACCTGTGAAGAGTCT	1620
Qy	1621	aaatggaactcgccttcaagtgagctacaagtatctttgtccaaggttagagagcctt	1680
Db	1621	AAATGGAATCTGCCCTTCCAGGTGGCTACAAGAGTATCTTTGTTCAGGTAGGAGACCTT	1680
Qy	1681	gtggcctccaagtgcacttccagggcctgcttggccctctcttcacgggtctgctcagtt	1740
Db	1681	GTGGCCTCCAGCTGCACCTTCCAGGGCGCTGCTTGGGCCCTCTTCTACGGGCTCTGTCTCAGT	1740
Qy	1741	ctctatgaatccttcagggaagattcatatttagactcttcaacagtttgactgaattt	1800
Db	1741	CTTCTATGAATCCTTCAGGGCAGATTCATATTTAGACTCTTTCACAGTTTGACCTGAGTGT	1800
Qy	1801	tggccagaataaagtgacatttagttgttgcttgaaggatgacttaaaattattagaca	1860
Db	1801	TGGCCAGAAATAGGTGCACATTTAGTTGTGTGCTGTAGTGATGACTTAAATATTTAGACA	1860
Qy	1861	tggtgtgaggcctgcatctcactcttgctcttttttttggccctccagtgctttgggt	1920
Db	1861	TGGTGTGTAGGCGTCATTTCCCTACTCTTGCCCTTTTTTTTTTTTGGCCCTCCAGTGTGTTGGGT	1920
Qy	1921	agtttggctccctcagccaaggaacagagaagttggagctctggagctggagctacat	1980
Db	1921	AGTTTGTCTCCCTCAGGCCAAAGGCAACAGAGAAGTTGGAGGTCTGGAGTGGCTACAT	1980
Qy	1981	aattttacagctgcaattctctggctgcacttcacaaatgtatacaaaactaaatacaa	2040
Db	1981	AATTTTACAGGACTGCAATTTCTCTGCTGCATTCACAAATGTATACAACTAAATACAA	2040
Qy	2041	gtcctgtgtttttatcaaggaggctgatcaataataataaataaaggggctgggtc	2100
Db	2041	GTCTCTGTGTTTTATCACAGGAGGCTGATCAATATAATGAATTAAGAAATTAAGGGGCTGGTC	2100





Db 28640 CTCACCTGCTTCTGCTGAGTGACTGAACATCAATAACAGAGCCCGGACCGGCGGG 28699

Qy 2804 gaggagagagacacagggctttgaccgatagtagtaacctctgcgtcgtgtagccgaatct 2863

Db 28700 GAGGAGGAGAGACACAGGCTTTGACCGGATAGTAACCTCTCGCTCGGTGACGCGGAATCT 28759

Qy 2864 ataaaggaactagtcacggcaaaaccccgtaattgcagcagagagtagtgaggccgg 2923

Db 28760 ATAAAGGAACCTAGTCCCGCAAAACCCCGTAATTGCGAGCGAGAGTAGTGAGGGCCGG 28819

Qy 2924 gaccgcagagcagcagccacctctctcccggtgctgcgagcagcagcagcagcagcagc 2983

Db 28820 GACCGCAGAGCGAGCGACCCCTTCTCCCGGCTGCGGACGCGAGCGGGGAGCT 28879

Qy 2984 ccgcgcacaaacagagcgggtctccagggcgtttgctccttctttttcccggttctg 3043

Db 28880 CCGCGCACCAACAGAGCGGCTTCTCAGGGCGCTTCTCTCTTTTTCCTCCGCTCTG 28939

Qy 3044 ttctccctctccggaaggtgtgcaagggttaggagaaagagacgacacacacaaa 3103

Db 28940 TTTTCTCCCTCTTCCGGAAGGCTTGTCAAGGGTAGGAGAAAGAGACGACACACAAA 28999

Qy 3104 gtgaaacacaggttaagaggtctccagtgacttacttggcgttatgtttgttctcgag 3163

Db 29000 GTGAAACACAGGTAAGAGGCTCTCCAGTGACTTACTTGGGCGTTATTTGTTTTCGAG 29059

Qy 3164 gcaagagccttcggaagtgtcgggttcggggactttgatccgagccccacacccc 3223

Db 29060 GCCAAGGAGGCTTCGGGAAGTGTGCGGTTCGGGGACTTTGATCCGAGCCACATCCC 29119

Qy 3224 caccactt 3231

Db 29120 CACCATT 29127

RESULT 3

AL359182/c

LOCUS

DEFINITION

Human DNA sequence from clone RP11-217B7 on chromosome 9, complete sequence.

ACCESSION

AL359182

VERSION

AL359182.20

GI:18151453

KEYWORDS

HTG.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 96717)

Direct Submission

Submitted (11-JAN-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

On Jan 15, 2002 this sequence version replaced gi:18121468.

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C.elegans/wormpep

This sequence was generated from part of bacterial clone contigs of human

chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr9

RP11-217B7 is from the library RPCI-11.1 constructed by the group of Pieter de Jong. For further details see

http://www.chori.org/bacpac/home.htm

VECTOR: pBACE3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-217B7. It may be shorter because we sequence overlapping sections only once, except for a short overlap.

The true right end of clone RP11-217B7 is at 96717 in this sequence. The true left end of clone RP11-122F10 is at 72980 in this sequence. The true right end of clone RP11-31J20 is at 2000 in this sequence.

FEATURES

source

1..96717

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="9"

/clone="RP11-217B7"

/clone\_lib="RPCI-11.1"

3238..3278

/note="Sequence from AF275948 sequenced by National Institutes of Health, National Heart, Lung and Blood Institute, Bethesda, MD 20892, USA."

84249..84273

/note="Sequence from overlapping clone RP11-122F10 (AC026643). Assembly confirmed by restriction digest."

92050..92163

/note="Sequence from reads from a short insert library derived from a single pUC clone. Restriction digest data confirm the assembly."

92411..92557

/note="Sequence from reads from a short insert library derived from a single pUC clone. Restriction digest data confirm the assembly."

BASE COUNT 27673 a 21138 c 20380 g 27526 t

ORIGIN

Query Match 96.6%; Score 3121.6; DB 9; Length 96717;

Best Local Similarity 98.8%; Pred. No. 0;

Matches 3207; Conservative 0; Mismatches 9; Indels 30; Gaps 5;

Qy 2 caggcgatgctgcaggtccctgtaattcagttactcgggaggtgaggttgcaatgag 61

Db 20151 CCGGGCATGCTGGCAGTGCCTGTATCTCAGCTACTCGGAGGTGGAGTTGCAATGAG 20092

Qy 62 ccagatcgccacattgctactccagcctgggcaacaaagggtgaaactccatctcaatta 121

Db 20091 CCCAGATCGCACATTGCACTCCAGCTGGGCAACAAAAGGTGAAACTCCATCTCAATTA 20032

Qy 122 aaaaaaaagaatgatttgggtggtcgaattcgaatagtagtagaagaagagagagga 181

Db 20031 AAAAAAAATAATGATTTGGTGGTGCAGTTCATAATAGTAGGAGAAAGGAGAGAGA 19972

Qy 182 gatggaggtcgaggagatcttaattactctctaaatcatgctaggaagataaacacct 241

Db 19971 GATGGAGGTGAGGAGAGATCTAATTACTCTCTAAATCATGCTAGGAAAGATAACACCTT 19912

Qy 242 ttaataacactctctcttttataacacatctctgcaggagctcaagggttcaacaa 301

Db 19911 TTAATAACACTCTCTCTCTTTTATAACATCATTTCTGCAAGGAGCTCAAAAGGTTCAACAA 19852

Qy 302 agtcaatttcagaaacccctttgaggagacagacatacatctctctccattttaa 361

Db 19851 AGTTCACTTTTCAGAAAACCCCTTTGAGGAAGACAGAAATATACATCTCTCTCCATTTAA 19792

Qy 362 agatgaagaacagccgggcacaaatggctaaatgcctgtaatccccagcactttgggagcc 421

Db 19791 AGATGAAGAACAAGCCGGGCACAAATGGCTAATGCTCTGTAAATCCAGCAGCTTTGGGAGGC 19732

Qy 422 tgaggccagaggatcgtcttgagctccagagtttgagaccagcctgggataacatggcaaaa 481











QY 602 tgagccatgattgatacactgcactccagctggtggtgagacagcaagaccctgtctcaa 661  
 Db 31667 TGAGCCATGATTGATGACTGCACTCCAGCTGGGTAGACAGACGAAGACCTGTCTCAA 31726  
 QY 662 aaaaaagaaatgaagagaaagaaagagagagagagagagagagagagagagag 721  
 Db 31727 AAAAAAGAAATGAAGAGAAAGAAAGAAAGAGAGAGAGAGAGAGATGAGGGAGGAGG 31786  
 QY 722 aggggggaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaag 781  
 Db 31787 AGGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 31846  
 QY 782 atgaacagagggag 841  
 Db 31847 ATGAACAGAGGCAGAGAGACTTTACGTAAATTGCTCATCATGTGTTCTCAAGTTGAC 31906  
 QY 842 ccaaaacccaattttatgacaaggtttatcttttgaactgagggaggggtccctctc 901  
 Db 31907 CCCAAAACCCAAATTTATTGACCAAGGTTATTCTTTGACTGAGGCAAGGGGGTCCGCTTC 31966  
 QY 902 ctgggcttgggcttttagaaagctcatctctgcttctgagatccatccctttttt 961  
 Db 31967 CTGGGCTTGGGCTTTAGAAAGCTCATCTCTGGCCCTTTCTGAGATCCATCCCTTTCTTTT 32026  
 QY 962 tattttcttgacacggagcttctgctctgctcactcaggtgagtgagtgcatgatct 1021  
 Db 32027 TATTTTCTTGACACGAGCTCTGCTCTGCTACTCAGCTGAGTGCACTGSCATGATCT 32086  
 QY 1022 cgaactcagttaacctctgctcccggttcaagggattctcctgctcagcctcctgag 1081  
 Db 32087 CGACTCAGTGTAACTCTGCTCCCGGTTCAAGCGATTCTCCTGCTCAGCCTCCTGAG 32146  
 QY 1082 ataacagggccgcacacacatctgctgctgctgctgctgctgctgctgctgct 1141  
 Db 32147 ATAACAGGGCCCGCCACACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 32206  
 QY 1142 catcatgttggccaggttgggttgcagctcctgagctgagtgagtgagtgagtgag 1201  
 Db 32207 CATCATGTTGGCCAGCTGTTGTTGCAACTCTGCACTGAGGTGAGTGCTGCCACTTGGCC 32266  
 QY 1202 tcccaagtctgggattacagcatgagccactgagccactgagccactgagccacttcc 1261  
 Db 32267 TCCCAAGTCTGGGATTACAGCATGAGCCACTGCGCCAGCTCAGATCCATCCCTTTC 32326  
 QY 1262 taaggcaaacagtcctagtgcaagggccatgccacccagagttatgagtaactggg 1321  
 Db 32327 TAAGGCAAACAGTCCATGGTGCAAGGGGCCATGCCACCCAGAGTTATGATACCTGGG 32386  
 QY 1322 actccagaattccttgcctgggtccctccacatgcaactccagggcctgttgggctct 1381  
 Db 32387 ACTCCAGAATTCCTTGCCTGGTGGCTCCACATGCATTCAGAGGCCCTGCTTGGGCCCT 32446  
 QY 1382 tctatgctctgctgagtgatgtagaacactgagtgagtagctgagtgagtgagtgag 1441  
 Db 32447 TCTATGCGTCTGTCTGCTGAGTGTGTATAGAACCACTGATGTAGTACCTGGGCTTGAGCG 32506  
 QY 1442 tggcctgagatccttctgactgagcatgagggggttggagctgaaatgctgcat 1501  
 Db 32507 TGCCCTTGGAGATCCTGTTGACTGTAGCATGGAGGGGCTTGTGACGCTGAATGTCTGAT 32566  
 QY 1502 gcagggtgtggagttctggaatagtaggagctgaggtggagagagagagagagagag 1561  
 Db 32567 GCAGGTGGTGGGATGCTGATATGATGAGCTGGAGTGGGAGAGAGAGTAGGCTTGG 32626  
 QY 1562 ggcagctctcatgccaactcattcttgcccaaaactcaggtcaaacctgtgaagagtcta 1621  
 Db 32627 GGCAGCTCTCTCATGCACTTATCTGGCCAAAACCTCAGGTCAAACTGTGAAGAGTCTA 32686  
 QY 1622 aatgtgaatcctccctcaagtggtactcaaaaggtatctttgtcaaggttagagaccttg 1681  
 Db 32687 AATGTGAATCTGCCCTTCAAGGTGGCTACAAAGGTATCTTTGTCAAGGTAGGAGACCTTG 32746

QY 1682 tggcctcacgtgacttccagggcctgtgggctcttctcagggctctgtcctgagtc 1741  
 Db 32747 TGGCCTCCACGTGCACTTCCAGGGCCTGCTTGGGCTCTTCTACGGGCTCTGCTCTGAGTC 32806  
 QY 1742 ttctatgaat----ccttcagggcagattcatatttagactcttcacagtttgactgag 1797  
 Db 32807 TTCTATGAATCTGCCCTTTCAGGGCAGATTCATATTAGACTCTTTCACAGGGGATCAATA 32866  
 QY 1798 ttgtggccagaataaggtgacatttagttgttgctgagtgagtgacttaaatatttag 1857  
 Db 32867 GGTNGCCCATGCACTGAGCTCAGT-----NNNAG 32895  
 QY 1858 acatggtgtgagggcctgcattcctactctgtccttttttttttttttttttttttttt 1917  
 Db 32896 NCCCCCTCTCCNN 32955  
 QY 1918 ggtagttttgtcctccctacagccaaaggaaggaaggaaggaaggttgagtgagtgag 1977  
 Db 32956 GGTAGTTTGTGCTCCCTACAGCCAAAGGCAACAGAGAAAGTTGGAGGTCTGGAGTGGCTA 33015  
 QY 1978 cataattttacagactgaattctctgctgacttcacaaaatgatacaaaaataaata 2037  
 Db 33016 CATAATTTTACAGACTGCAATTTCTGGCTGCATTTACAAATGTATACAAAGTAAATA 33075  
 QY 2038 caagtctctgtttttatcacagggaggtgatcaataataataataataataataataata 2097  
 Db 33076 CAAGTCTCTGTTTATCAGAGGAGGCTGATCAATATATATATATATATATATATATAT 33135  
 QY 2098 gtccatattgttctgttt 2157  
 Db 33136 GTCCATATTGTTCTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT 33194  
 QY 2158 gctcctcctcctcaatttatgaagagaagcagtaagaatgtctcctcctcctcctgag 2217  
 Db 33195 GCCCCTCCCTCTCAATTTATGAGAGAAACACCTGAGGGGTCTCTTTTGGCCCTTAAA 33254  
 QY 2218 ggacctggggagctcagctgggaatctccaagggcagtagtgctcctcctcctcctcct 2277  
 Db 33255 AGAAGTGTGCTTAACTCGGGAATTTCCAGGCGTAAATTCCTCTCATATATATATATAT 33314  
 QY 2278 gtccaggtttgtgggggaaacaaacagcagccattaccagagagactgtccgctctcc 2337  
 Db 33315 AACAGGGTTGATCACAAAAACAAAGCAGCCATTTACCCAGAGGACTGTCTGCTCTCC 33374  
 QY 2338 cctcaacccagccttagcctttgaaaggaacaaagaaacaaagaaacaaagaaatgattggc 2397  
 Db 33375 CCTCACCCAGCCTTAGGCTTTTGAAGGAAACAAAGAACAAAGAAATGATGGCGTCC 33434  
 QY 2398 tgaggagattcagcctagagctctctcctcccaatcctcctcctcctcctcctcctcct 2457  
 Db 33435 TGAGGGAGATTCAGCCTAGAGCTCTCTCTCCCCCAATCCCTCCTCCCTGAGAAACT 33494  
 QY 2458 aacaaagaaaaaaattgcggaaagcagagatttagaggagcaaatccactgggtgccc 2517  
 Db 33495 AACAAAGAAAAAAATTTGCGAAAGCAGCATTTAGAGGAAGCAAAATTCACACTGGTGCC 33554  
 QY 2518 cttggctccggggaacgttgagactagagagctgcggcgagcccccagagccagcgtctcc 2577  
 Db 33555 CTTGGCTGCCGGGAACGTGGACTAGAGAGTCTGCGGCGCAGCCCGCAGCCGAGCGCTTCC 33614  
 QY 2578 cgcgcgtcttaggc 2637  
 Db 33615 CGCGGCTTTAGGCTTGGCGGCGCCCGGGGGGGAAGGCGCAGCAGACCGCGGAGCCCTTAAG 33674  
 QY 2638 aacctgtgtacactccaccc 2697  
 Db 33675 ACACCTGTGTACCTCCACCCACCCACCCACCCACCCACCCACCCACCCACCCACCTAGAT 33734  
 QY 2698 gt 2757  
 Db 33735 GT 33794  
 QY 2758 tgctgagtgactgaactacataaaacagagggcggaagggggggggggggggggagagca 2817









RESULT 7  
 LOCUS AX060715 1643 bp DNA linear PAT 22-JAN-2001  
 DEFINITION Sequence 3 from Patent WO0078972.  
 ACCESSION AX060715  
 VERSION AX060715.1 GI:12406104  
 KEYWORDS human.  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1643)  
 AUTHORS Lawn,R.M., Wade,D. and Garvin,M.  
 TITLE Regulation with binding cassette transporter protein abcl  
 JOURNAL Patent: WO 0078972-A 3 28-DEC-2000;  
 CV THERAPEUTICS, INC. (US)  
 FEATURES Location/Qualifiers  
 source 1..1643  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 BASE COUNT 370 a 413 c 457 g 403 t  
 ORIGIN  
 Query Match 46.08; Score 1485.4; DB 6; Length 1643;  
 Best Local Similarity 98.18; Pred. No. 0;  
 Matches 1615; Conservative 0; Mismatches 6; Indels 26; Gaps 10;  
 QY 1345 gcctccacatgcactccaggcctctgtggcctctctatgcgtctgctcgttgactg 1404  
 Db 16 GGCCTCCACATGCACCTCCAGGCCCTGCTGG - CTCCTCATGGGTCTGCTGAGTGT 73  
 QY 1405 gatagaaccactgatgtgagtacctggccttgagccctggcctggagatcctgttgactg 1464  
 Db 74 GATAGAACCACTGATGTGAGTACCTGGGCTTGAG - CGTGGCCTGGAGATCGTGTGACTG 132  
 QY 1465 tagcatggaggggcttgcagctgaatctctcatgcaggtgggtggagttctggaat 1524  
 Db 133 TAGCATGGAGGGGCTTGT - CAGCTGAATGCTGTATGACAGGTGGTGGGAGTCTGGAAT 191  
 QY 1525 atgatgagctggaggtgggaagagaagtaggtgtgggcagctctctatgccacctca 1584  
 Db 192 ATGATGAGCTGGAGGTGGGAAGAGAAAGTAGAGCTTGGGGCAGCTCTCTCATGCCACCTCA 251  
 QY 1585 ttctggccaaactccaggtcaaacctgtgaagagctctaaatgtaaatcgcctctcaaggt 1644  
 Db 252 TTCTGGCCAAACTCAGGTCAACTGTGAAGAGTCTAAATGTGAATCTGCCCTTCAAGGT 311  
 QY 1645 ggctacaaaggtatctttgttcaaggttaggagacctgtggcctccacgtgcacttccagg 1704  
 Db 312 GGCTACAAAGGTATCTTTGTCAAGGTAGGAGACCTTGTGGCTCCACGTGCACCTCCAGG 371  
 QY 1705 gcctgtgtggcctcttcacgggtctgtcctgagttcttatgaat --- ccttcaagggc 1761  
 Db 372 GCCGTCTT - GCCCTCTTCTACGGGTCTGCTGAGTCTTCTATGAATCTCCCTTCAGGGC 430  
 QY 1762 agattcatattagactcttcacagttttgacctgagttttggccagataaaggtgacatt 1821  
 Db 431 AGATTCAATATTAGACTCTTCACAGTTTGGACCTTGGCCAGAAATAAGGTGACATT 490  
 QY 1822 tagttgttggttgatggtatgacttaaaattattagac -- atgggtgtgtaggctgcaatt 1879  
 Db 491 TAGTTTGTGGCTGTGATGAATGACTTAAATATTATTAGACATATGTTGTGTAGGCTTGCAATT 550  
 QY 1880 cctactcttgccttttttttggccctccaggttttgggtagttttgctt - cccctacag 1938  
 Db 551 CCTACTCTTGCCCTTTTTTTTGGCCCTCCAGTGTTTTGGGTAGTTTGGTCTCCCTACAG 610  
 QY 1939 ccaaaggcaaacagagaggttgaggtctggagtggtgacctacataattttacacgactgcaa 1998  
 Db 611 CCAAGGCAACAGATAGTTTGGAGGTCTGGAGTGGCTACATATATTTTACACGACTGCAA 670

Qy	1999	tctctggctgcacttcacaaatgtatacaaacataacaagtctcgtgtttatacac	2058
Db	671	TTCTCTGGCTGCACATTCACAAATGTATACAACATAAATACAAGTCCCTGTTTTATCAC	730
Qy	2059	aggagagctcatcaataataaaaattaaaaagggcgctgg-tccatatgtctcgtgcttt	2117
Db	731	AGGAGGCGCTGATCAATATATAAATGAATTAAGAAGGGCGTGCTCCCATATATGTTCTGTGTTT	790
Qy	2118	tgttgttggtttgtttgttttctttttttgttttttggcctccttcctctcaattta	2177
Db	791	TTG-----TTTGTGTTCTTTTGTTTTGTGGCTCCTTCCTCTCAATTTA	841
Qy	2178	tgaagaagaacagtagaatgttctctcgggtcctctgaaggacctggggagctcaggct	2237
Db	842	TGAAGAAGACAGTAGAATGTTCTCTCGGGTCCTCTGAGGAGACCCTGGGGAGCTCAGSCT	901
Qy	2238	gggaatcccaagcgagtagtgcgctatcaaaaatcaaaagtcacaggtttgtggggggaa	2297
Db	902	GGGAATCTCCAAGCGAGTAGTGCCTCATCAAAAATCAAAGTCCAGGTTTGTGGGGGGAA	961
Qy	2298	aacaaaagcagcccattaccagaggaactgtccgccttccccctacccccagcctaggcct	2357
Db	962	AACAAAAGCAGCCATTACCCAGAGACTGTCCGCCTTCCCCTCACCCACAGCTTAGGCCT	1021
Qy	2358	tgaagaagaacaaaagacaagacaaaaatgatggcgctcgtgagggagatcagccctaga	2417
Db	1022	TTGAAGAGAAACAAAAACAAGACAAAATGATGTGGCTCTTGAGGGAGATTCACGCCTAGA	1081
Qy	2418	qctctctcccccaatccctccctccgctgagaaactaacaaaggaaaaaaaattg	2477
Db	1082	GCTCTCTCTCCCCCAATVCCCTCCTCCGCGCTGAGGAACACTACAAAGGAAAAANAATTG	1141
Qy	2478	cggaaagcaggatttagaggaagcaaatccactggtgcccttggctgcggggaactgg	2537
Db	1142	CGGAAGACAGGATTTAGAGGAAGCAAAATCCACTGCTGGCTTGCTCGCGGAACGCTGG	1201
Qy	2538	actagagagtctcggcgacagccccagagcccaagcgtccccgcgogtcttagccgcggcg	2597
Db	1202	ACTAGAGAGTCTCGCGCGACGCCCGGAGCCACAGCTTCCC CGCGGTCTTAGCGCGGCGG	1261
Qy	2598	qcccgcggggggaaagggaacgagaccgcggacccctaagacacctgctgacctccac	2657
Db	1262	GCCGCGCGGGGANGGGNACGCAGACC GCGACCCCTAAGACACCTGCTGTACCTCCAC	1321
Qy	2658	ccccccccccccccccactccccccccctcctagatgtgtcgtggcgctgaacg	2717
Db	1322	---- -CCCCACCCACCCACCTCCCCCCAACTCCCTAGATGTGTCTGGCGGCGGTGAACG	1376
Qy	2718	tcgcccgttaaaggcgcgggccccgcgtccacgtcttctgtgtagtactgaactaca	2777
Db	1377	TCGCCCGTTTAAGGGCGGGCCCCCGCTCCACGTGCTTCTGTGTGAGTGACTGAACCTACA	1436
Qy	2778	taacagagcgcggaaggggcggggagagagagagagagagagagagagagagagagag	2837
Db	1437	TAAACAGAGCGCGGAACCGGGCGGGAGGAGGAGACAGGCTTTTGACCGATAGTAA	1496
Qy	2838	ccctcgcgtcgtgcagccgaatctaataaagggaactagtcctccggcaaaaaaccccgtaa	2897
Db	1497	CCCTCGCGCTCGGTGCAGCCGAATCTATATAAAGGAACCTAGTCCCGGCAAAAACCCCGTAA	1556
Qy	2898	tgcgagcgagagtgtgtggggcgggaccccgagacgagcgagccgacctctctcccg	2957
Db	1557	TTGCGAGCGAGAGTAGTGGGCGGGACCCGACAGCGCGAGCCGACTTCTCTCCCGG	1616
Qy	2958	gctcggcgaggggcaggggcgggagctc	2984
Db	1617	GCTCGCGCAGGGCAGGCGGGAGCTC	1643

RESULT 8  
AX060894  
LOCUS

1643 bp DNA linear PAT 22-JAN-2001





```

SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 175064)
REFERENCE
AUTHORS
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE
Homo sapiens, clone RP11-1M10
JOURNAL
Unpublished
REFERENCE
AUTHORS
1 (bases 1 to 175064)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhgalter,B.,
Brooke,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., DeArelano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (21-Oct-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 22, 2000 this sequence version replaced gi:6454033.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Web site: http://www-seq.wi.mit.edu
Center code: WIBR
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: 1JM10
Center clone name: 1JM10
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 117571 bases at least Q40
Consensus quality: 145749 bases at least Q30
Consensus quality: 160940 bases at least Q20
Insert size: 185000; agarose-fp
Insert size: 171264; sum-of-contigs
Quality coverage: 2.9 in Q20 bases; agarose-fp
Quality coverage: 3.2 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 39 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1003: contig of 1003 bp in length
* 1004 1103: gap of 100 bp
* 1104 2634: contig of 1531 bp in length
* 2635 2734: gap of 100 bp
* 2735 4415: contig of 1681 bp in length
* 4416 4515: gap of 100 bp
* 4516 5785: contig of 1270 bp in length
* 5786 5885: gap of 100 bp
* 5886 7879: contig of 1994 bp in length
* 7880 7979: gap of 100 bp
* 7980 9686: contig of 1707 bp in length
* 9687 9786: gap of 100 bp
* 9787 12253: contig of 2467 bp in length
* 12254 12353: gap of 100 bp
* 12354 15228: contig of 2875 bp in length
*
15229 15328: gap of 100 bp
15329 17200: contig of 1872 bp in length
17201 17300: gap of 100 bp
17301 20131: contig of 2831 bp in length
20132 20231: gap of 100 bp
20232 22587: contig of 2356 bp in length
22588 22687: gap of 100 bp
22688 25707: contig of 3020 bp in length
25708 25807: gap of 100 bp
25808 28184: contig of 2377 bp in length
28185 28284: gap of 100 bp
28285 31338: contig of 3054 bp in length
31339 31438: gap of 100 bp
31439 34299: contig of 2861 bp in length
34300 34399: gap of 100 bp
34400 38318: contig of 3919 bp in length
38319 38418: gap of 100 bp
38419 42835: contig of 4417 bp in length
42836 42935: gap of 100 bp
42936 45448: contig of 2513 bp in length
45449 45548: gap of 100 bp
45549 48116: contig of 2568 bp in length
48117 48216: gap of 100 bp
48217 52618: contig of 4402 bp in length
52619 52718: gap of 100 bp
52719 56592: contig of 3874 bp in length
56593 56692: gap of 100 bp
56693 59635: contig of 2943 bp in length
59636 59735: gap of 100 bp
59736 63661: contig of 3926 bp in length
63662 63761: gap of 100 bp
63762 68437: contig of 4676 bp in length
68438 68537: gap of 100 bp
68538 71458: contig of 2921 bp in length
71459 71558: gap of 100 bp
71559 76888: contig of 5330 bp in length
76889 76988: gap of 100 bp
76989 82113: contig of 5125 bp in length
82114 82213: gap of 100 bp
82214 88220: contig of 6007 bp in length
88221 88320: gap of 100 bp
88321 93499: contig of 5179 bp in length
93500 93599: gap of 100 bp
93600 97901: contig of 4302 bp in length
97902 98001: gap of 100 bp
98002 103016: contig of 5015 bp in length
103017 103116: gap of 100 bp
103117 109178: contig of 6062 bp in length
109179 109278: gap of 100 bp
109279 117307: contig of 8029 bp in length
117308 117407: gap of 100 bp
117408 124079: contig of 6672 bp in length
124080 124179: gap of 100 bp
124180 131281: contig of 7102 bp in length
131282 131381: gap of 100 bp
131382 138059: contig of 6678 bp in length
138060 138159: gap of 100 bp
138160 145491: contig of 7332 bp in length
145492 145591: gap of 100 bp
145592 157391: contig of 11800 bp in length
157392 157491: gap of 100 bp
157492 175064: contig of 17573 bp in length.
-----
FEATURES
source
1. 175064
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="rp11-1m10"
/clone_lib="RPC1-11 Human Male BAC"
/misc_feature
1. 1003
/note="assembly_fragment"
misc_feature
1104. 2634
/note="assembly_fragment"
misc_feature
2735. 4415

```



```

/misc_feature /note="assembly_fragment"
4516. .5785
/misc_feature /note="assembly_fragment"
5886. .7879
/misc_feature /note="assembly_fragment"
7980. .9686
/misc_feature /note="assembly_fragment"
9787. .12253
/misc_feature /note="assembly_fragment"
12354. .15228
/misc_feature /note="assembly_fragment"
15329. .17200
/misc_feature /note="assembly_fragment"
17301. .20131
/misc_feature /note="assembly_fragment"
20232. .22587
/misc_feature /note="assembly_fragment"
22688. .25707
/misc_feature /note="assembly_fragment"
25808. .28184
/misc_feature /note="assembly_fragment"
28285. .31338
/misc_feature /note="assembly_fragment"
31439. .34299
/misc_feature /note="assembly_fragment"
34400. .38318
/misc_feature /note="assembly_fragment"
38419. .42835
/misc_feature /note="assembly_fragment"
42936. .45448
/misc_feature /note="assembly_fragment"
45549. .48116
/misc_feature /note="assembly_fragment"
48217. .52618
/misc_feature /note="assembly_fragment"
52719. .56592
/misc_feature /note="assembly_fragment"
56693. .59635
/misc_feature /note="assembly_fragment"
59736. .63661
/misc_feature /note="assembly_fragment"
63762. .68437
/misc_feature /note="assembly_fragment"
68538. .71458
/misc_feature /note="assembly_fragment
clone_end:SP6
vector_side:left"
71559. .76888

Query Match 38.0%; Score 1227.8; DB 2; Length 175064;
Best Local Similarity 98.7%; Pred. No. 8:2e-311;
Matches 1263; Conservative 0; Mismatches 2; Indels 14; Gaps 2;

QY 1953 agaagttgaggtctgagtggtggtacataattttacacgactgcaattctctgtgtgcac 2012
DB 2735 ATAAGTTGGAGTCTGGAGTGGCTACATAATTTTACAGCAGCTGCAATTCCTGCTGCAC 2794
QY 2013 ttcacaaatgtatacaaaactaaatacaagtcctgtgtttttatcacaggaggtgatca 2072
DB 2795 TTCACAAATGTATACAACTAAATACAAAGTCTGTTTATCACAGGAGGCTGATCA 2854
QY 2073 atataatgaataaaaggggctggtccatattgttctgtgttttttttttttttttttt 2132
DB 2855 ATATAAATAAATTAAGAGGGGGCTGGTCCATATTTGTTGTTGTTTGTGTTTGTGTTT 2905
QY 2133 gttgtttcttttttttttttttttttttttttttttttttttttttttttttttttttt 2192
DB 2906 GTTTGTTTCTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 2965
QY 2193 agatgttctctcgtggtctcgtgaggaactcgtgaggaactcgtgaggaactcgtgagga 2252
DB 2966 AGATGTTCCTCTCGGGTCTCTGAGGACCTGGGGAGCTCAGGCTGGGGAATCTCCAAAGC 3025

```

```

QY 2253 agtaggtcgctctatcaaaaatcaaaagtcaggtttgtggggggaacaaacagaccaca 2312
DB 3026 AGTAGGTGCGCTATCAAAAATCAAAAGTCCAGGTTTGTGGGGGGAACAAACAGACGCCA 3085
QY 2313 ttacccagaggaactgtccgcttcctccctcaccacagccttaggcctttgaaagaaacaa 2372
DB 3086 TTACCCAGAGGAGTGTCCGCTTCCCTCACCACAGCCTAGCGCTTTGAAAGGAACAAA 3145
QY 2373 agacaagacaaaatgattggcgctcctgaggaagattcagcctagagctctctctccccc 2432
DB 3146 AGACAGACAAAATGATTGGCGTCTGAGGAGATTACAGCTTAGAGCTCTCTCTCCCCCA 3205
QY 2433 atccctccctccggtgaggaatacaaaagaaataaaattgcggaagcagattt 2492
DB 3206 ATCCCTCCCTCCGCTGAGGAAACTAACAAAGGAATAAATTTGCGGAAGCAGGATTT 3265
QY 2493 agaggaaagcaaatccacttggtgctcccttggtcccggaacgtggactagagagtgctgg 2552
DB 3266 AGAGGAAGCAAAATTCACACTGCTGCCCTTGGCTGCCGGAACGTGGACTAGAGAGTCTGCG 3325
QY 2553 ggcagcccccgagccccagcgcttcccgcggttcttaggcggcgccggcggggggaa 2612
DB 3326 GCGAGCCCCGAGCCGAGCGCTTCCGCGCGCTTTAGCGCGCGGCGCCGCGGGGGGAA 3385
QY 2613 ggggacgacagcgcgagaccctaagacacctgctgtacctccacccccaccccccca 2672
DB 3386 GGGAGCGACAGCCGCGGACCTAAGACACCTGCTGTACCTCCAC-----CCCCACCCCA 3440
QY 2673 cccactccccccaaatccctagatgtgctgtgggggctgaaactgcgcctgttaaggg 2732
DB 3441 CCCACCTCCCCCAACTCCCTAGATGTGCTGGCGGCTGAACGTCGCCCTTTAAGGG 3500
QY 2733 gcgggccccggctccacgctgtcttctgtagtgactgaactacataacacagagggcg 2792
DB 3501 GCGGGCCCCGGCTCCACGCTTTTCTGCTGAGTGAACCTCTGCGCTCGGTG 3560
QY 2793 aagggggcggggagggagagacacagcctttgacgtagtaacactctgcgcctcggtg 2852
DB 3561 AACGGGCGGGGAGGAGGAGACAGGCTTTGACCGATAGTAACCTCTGCGCTCGGTG 3620
QY 2853 cagccgaatctataaaggaaactagtcctcgcgcaaaacccccgtaattgagagcgagtg 2912
DB 3621 CAGCCGAATCTATAAAGGAACCTAGTCCCGGCAAAACCCCGTAATTGCGAGCGAGAGTG 3680
QY 2913 agtggggcggggacccgcagagccgagccgacccctctctcctcggcggtcgggcagggcag 2972
DB 3681 AGTGGGCGCGGACCCGCGAGCGCGACCCCTTCTCTCCCGGCTGCGGCGAGGCGAG 3740
QY 2973 ggcgggagctccgcacacacagacggcttctcagggcgcttgcctctgttttttttt 3032
DB 3741 GCGGGGAGCTCCGCGCACCAACAGAGCGGTTCTCAGGGCGCTTTGCTCTTGTGTTT 3800
QY 3033 ccccggttctgttttctcccttctcggaggtctgtcaagggttaggagaaagagacg 3092
DB 3801 CCCCGGTTCTGTTTCTCCCTTCTCCGGAAGCTTTGCAAGGGGTAGGAGAAGAGACG 3860
QY 3093 caaacacaaagtggaaaacaggttaagaggtctccagtgacttacttggggttattgt 3152
DB 3861 CAAACAAAAGTGAACACAGGTAAGAGGCTCTCCAGTGACTTACTTGGCGCTTATTGT 3920
QY 3153 ttgtttcagggcgaagggcttcgggaagtgcgttcggttcgggacttgcacggag 3212
DB 3921 TTTGTTTCAGGGCCAGGAGGCTTCGGGAAGTGTCTCGGTTTCGGGGACTTTGATCCGGAG 3980
QY 3213 cccacacatccccaccactt 3231
DB 3981 CCCCACATCCCCACCACCTT 3999

```

RESULT 10  
HSA252201  
LOCUS  
DEFINITION

HSA252201 Homo sapiens partial ABC-1 gene for ATP-binding cassette  
1167 bp DNA linear PRI 10-APR-2001

```

transporter-1, 5'UTR and promoter region.
ACCESSION AJ252201
VERSION AJ252201.1 GI:12053757
KEYWORDS ABC-1 gene; ATP-binding cassette transporter-1; promoter.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.
Porsch-Oezcuemez, M., Langmann, T., Helmerl, S., Borsukova, H.,
TITLE The zinc finger protein 202 (znf202) is a transcriptional repressor
of atp binding cassette transporter 1 (abc1) and abcg1 gene
expression and a modulator of cellular lipid efflux
J. Biol. Chem. 276 (15), 12427-12433 (2001)
JOURNAL 21192304
MEDLINE 2
REFERENCE 2 (bases 1 to 1167)
AUTHORS Porsch-Oezcuemez, M.K.
TITLE Direct Submission
JOURNAL Submitted (05-JAN-2000) Porsch-Oezcuemez M.K., Institute for
Clinical Chemistry, University of Regensburg,
Franz-Josef-Strauss-Allee 11, 93042 Regensburg, GERMANY
FEATURES
source
1..1167
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/map="9q22-31"
/cell_type="leukocyte"
gene
1..1167
/gene="ABC-1"
promoter
1..942
/function="ABC-1"
TATA_signal
896..900
/gene="ABC-1"
intron
1148..1167
/gene="ABC-1"
/function="cholesterol efflux regulatory protein"
BASE COUNT 278 a 315 c 327 g 247 t
ORIGIN
Query Match 35.0%; Score 1130.4; DB 9; Length 1167;
Best Local Similarity 98.7%; Pred. No. 9.3e-286;
Matches 1165; Conservative 0; Mismatches 1; Indels 14; Gaps 2;
QY 1955 aagtgaggtctggagtgctacataattttacagactgcaattctctggtgcaatt 2014
Db 2 AAGTTGGAGTCTGGAGTGGCTACATAATTTTACAGACTGCAATTTCTTGGCTGCACAT 61
QY 2015 cacaatgtatatacaataatacaagtcctgtgtttttatcacagggaggctgatcaat 2074
Db 62 CACAATGTATACAACTAAATACAAAGTCTGTGTTTATATACAGGAGGCTGATCAAT 121
QY 2075 ataatgaataataaagggtctgtgtccatattgtctgtgtttttgtttgtttgtt 2134
Db 122 ATAATGAATATAAAGGGGCTGGTCCATATTTTGTGTTTGTGTTTGTGTTTGTGTTTGT 172
QY 2135 ttgttttttttttttttttttttttttttttttttttttttttttttttttttttttt 2194
Db 173 TTGTTTCTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT 232
QY 2195 atgtctctcgtgtctctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2254
Db 233 ATGTTCTCTCGGTCCTCTGAGGACCTGGGAGCTCAGGCTGGGAATCTCCAGGCAG 292
QY 2255 tagtgcctatcaataataagtcacagtttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2314
Db 293 TAGTGCCTTATCAAAATCAAAAGTCCAGGTTTGTGGGGGAAACAAAGAGAGCCCAT 352
QY 2315 accaagagacttcgccttcctccaccagcctagcctttgaaaggaacaaag 2374
Db 1167

```

---

```

Db 353 ACCCAGAGGAGTGTCCGCTTCCCTCACCACCGAGCTTAGGCCTTTGAAAGGAAACAAAG 412
QY 2375 acaagacaaaatgattgcgtcctagagagattcagcctagagactctctctccccaat 2434
Db 413 ACAAGACAAAATGATTGGCGTCTGAGGAGATTACGCTAGAGCTCTCTCTCCCAAT 472
QY 2435 cctcctcctcgtcgtgagaaactaacaaaggaagaaataatgctggaagcagagatttag 2494
Db 473 CCTCTCCTCCGCTGAGGAACTAACAAAGGAAATAATGCGGAAAGCAGGATTAG 532
QY 2495 aggaagcaaaatccactagtcctcctcctcctcctcctcctcctcctcctcctcctcctc 2554
Db 533 AGGAAGCAAAATCCACTGGTGGCTTGGCTCCGGGACGTGGACTAGAGAGTCTGGGGC 592
QY 2555 gcagcccgagcccgagcgtctcccgcgctcttaggcgcgcgcgcgcgcgcgcgcgcgcgcgc 2614
Db 593 GCAGCCCGAGCCAGCGCTTCCCGCGCTTCTAGCGCGCGCGCGCGCGCGCGCGGGAAGG 652
QY 2615 ggaacagacccgagccctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 2674
Db 653 GGACGACAGACCGCGGACCTAAGACACCTGCTGTACCTCCAC-----CCCCACCCACCC 707
QY 2675 cactcccccccaactcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 2734
Db 708 CACCTCCCCCAACTCCCTAGATGTCTGGCGGCTGAACGTGCGCCCGCTTTAAGGGGC 767
QY 2735 gggcccgctccacgtcgtcttctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 2794
Db 768 GGGCCCGCGCTCCACGTGCTTTCTGCTGACTGAACACTACATAAACAGAGCGCGGAA 827
QY 2795 gggggcgaggagagagagagagagagagagagagagagagagagagagagagagagagagag 2854
Db 828 CGGGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 887
QY 2855 gccgaatcataaaagggaactagtcctccgcgaacaccccgtaattgcgagcgagagtgag 2914
Db 888 GCCGAATCTATAAAGAGAACTAGTCCCGGCAAAACCCCGTAATTAATTCGAGGAGAGTGAG 947
QY 2915 tggggcgaggagagagagagagagagagagagagagagagagagagagagagagagagagag 2974
Db 948 TGGGGCGGGAGCCCGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGG 1007
QY 2975 cgggagctccgcgacacacacacacacacacacacacacacacacacacacacacacacacac 3034
Db 1008 CGGGAGGCTCCCGGACCAACAGAGCCCGTCTCAGGCGGCTTTGCTCCTCTGTTGTTTTC 1067
QY 3035 ccggtctgttttctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 3094
Db 1068 CCGGTTCTGTTTCTTCCCTTCTCCGGAAGGCTTGTCAAGGGGTAGGAGAGAGAGAGAGCA 1127
QY 3095 aacacaaagtgaagaaacagtaagagagctctccagtgcac 3134
Db 1128 AACACAAAAGTGAAGAACAGGTAAAGAGGCTCTCCAGTGAC 1167

```

---

```

RESULT 11
AF258623S1
LOCUS Homo sapiens ATP binding cassette transporter 1 (ABCAL) gene,
DEFINITION promoter and exon 1.
ACCESSION AF258623
VERSION AF258623.2 GI:8677405
KEYWORDS
SOURCE i of 4
human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.
AUTHORS 1 (bases 1 to 1167)
Pullinger, C.R., Hakamata, H., Duchateau, P.N., Eng, C.,
Aouizerat, B.E., Fielding, C.J., and Kane, J.P.
Analysis of hABC1 gene 5' end: additional peptide sequence,
promoter region, and four polymorphisms

```

JOURNAL REFERENCE	Biochem. Biophys. Res. Commun.	271	(2000)	In press
AUTHORS	2	(bases 224 to 1167)		
TITLE	Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C., Aouizerat,B.E., Fielding,C.J. and Kane,J.P.			
JOURNAL	Direct Submission			
REFERENCE	Submitted (20-APR-2000) Cardiovascular Research Institute, University of California, San Francisco, 505 Parnassus Avenue, San Francisco, CA 94143-0130, USA			
AUTHORS	3	(bases 1 to 1167)		
TITLE	Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C., Aouizerat,B.E., Fielding,C.J. and Kane,J.P.			
JOURNAL	Direct Submission			
REMARK	Submitted (23-JUN-2000) Cardiovascular Research Institute, University of California, San Francisco, 505 Parnassus Avenue, San Francisco, CA 94143-0130, USA			
COMMENT	Sequence update by submitter			
FEATURES	On Jun 23, 2000 this sequence version replaced gi:7769713. Location/Qualifiers			
source	1..1167			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/chromosome="9"			
	/map="9q31"			
	224..844			
promoter	/gene="ABCA1"			
exon	845..1147			
	/number=1			
BASE COUNT	278 a 313 c 328 g 248 t			
ORIGIN				
	Query Match 35.0%; Score 1130.4; DB 9; Length 1167;			
	Best Local Similarity 98.7%; Pred. No. 9,3e-286;			
	Matches 1165; Conservative 0; Mismatches 1; Indels 14; Gaps 2;			
QY	1955	aagttggaggtctggaagtgcacataatttacacgactgcaattctctgctgcacatt	2014	
Db	2	AAGTTGGAGGCTGGAGTGGCTACATAATTTACACGACTGCAAATTCCTGGCTGCACATT	61	
QY	2015	cacaatatgatacacaactaaataacaagtcctgtgttttatcacaggagcgtagtacaat	2074	
Db	62	CACAAATGTATACAAACTAAATACAAGTCCCTGTGTTTTATCACAGGAGGCTGATCAAT	121	
QY	2075	ataatgaaataaaaaggggctgtccattatgttgtctgtgtttgtttgtttgt	2134	
Db	122	ATRAATGAATTAAGAAGGGGCTGGTCCATATGTCTGTGTTTG-----TTGTG	172	
QY	2135	tgtttctttttttgtttgttgacctctctctctctcaattatgaagaagcagtaag	2194	
Db	173	TTGTTTCTTTTGTGTTTGTGGCTCCTCTCTCAATTTATGAAGAAGCAGTAAG	232	
QY	2195	atgttctctcggctcctcaggacctgggaacctcagctcagctgggaatctccaagcag	2254	
Db	233	ATGTTCTCTCGGTCCTCTGAGGACCCTGGGGAGCTCAGGCTGGGAATCTCCAAGGCAG	292	
QY	2255	tagtgcgctatacaaatcaaaagtcaggtttgtgggggaaacaaacagcagcccatt	2314	
Db	293	TAGTGCCTATCAAAAATCAAAGTCCAGGTTGTGGGGGAAACAAACAGCCCCATT	352	
QY	2315	accgagagactgtccgcttccccaccagacctgagccctttgaaggaacaaaaag	2374	
Db	353	ACCAGAGAGACTGTCCGCCCTCCCTCACCCCAGCTAGGCTGGGAATCTCCAAGGCAG	412	
QY	2375	acaagacaaaatgatggcgctcctaggagattcagctcagctcctctcccccaat	2434	
Db	413	ACAAGACAAAATGATTGGCGTCTCAGGGAGATTACAGCTAGAGCTCTCTCCCCCCAAT	472	
QY	2435	ccctccctccgctgaggaactaacaaaggaaaaaaaattgcggaagcaggatttag	2494	
Db	473	CCCTCCCTCCGGCTGAGGAACATAACAAGGAAAAAAAATTGCGAAGCAGGATTTAG	532	
QY	2495	aggaagcaaatccactggtgccttggctgccgggaacgtggactagagatgtcgggc	2554	

Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,  
Zimmer,A. and Zody,M.

# TITLE

## JOURNAL

### COMMENT

Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jul 13, 2000 this sequence version replaced gi:6705871.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L2512

Center clone name: L\_N\_10

-----

\* NOTE: This record contains 73 individual

\* sequencing reads that have not been assembled into

\* contigs. Runs of N are used to separate the reads

\* and the order in which they appear is completely

\* arbitrary. Low-pass sequence sampling is useful for

\* identifying clones that may be gene-rich and allows

\* overlap relationships among clones to be deduced.

\* However, it should not be assumed that this clone

\* will be sequenced to completion. In the event that

\* the record is updated, the accession number will

\* be preserved.

\* l

\* 872 971: gap of 100 bp

\* 972 1834: contig of 863 bp in length

\* 1835 1934: gap of 100 bp

\* 1935 2804: contig of 870 bp in length

\* 2805 2904: gap of 100 bp

\* 2905 3745: contig of 841 bp in length

\* 3746 3845: gap of 100 bp

\* 3846 4696: contig of 851 bp in length

\* 4697 4796: gap of 100 bp

\* 4797 5640: contig of 844 bp in length

\* 5641 5740: gap of 100 bp

\* 5741 6540: contig of 800 bp in length

\* 6541 6640: gap of 100 bp

\* 6641 7509: contig of 869 bp in length

\* 7510 7609: gap of 100 bp

\* 7610 8479: contig of 870 bp in length

\* 8480 8579: gap of 100 bp

\* 8580 9430: contig of 851 bp in length

\* 9431 9530: gap of 100 bp

\* 9531 10376: contig of 846 bp in length

\* 10377 10476: gap of 100 bp

\* 10477 11322: contig of 846 bp in length

\* 11323 11422: gap of 100 bp

\* 11423 12302: contig of 880 bp in length

\* 12303 12402: gap of 100 bp

\* 12403 13280: contig of 878 bp in length

\* 13281 13380: gap of 100 bp

\* 13381 14241: contig of 861 bp in length

\* 14242 14341: gap of 100 bp

\* 14342 15196: contig of 855 bp in length

\* 15197 15296: gap of 100 bp

\* 15297 16123: contig of 827 bp in length

\* 16124 16223: gap of 100 bp

\* 16224 17072: contig of 849 bp in length

\* 17073 17172: gap of 100 bp

\* 17173 18041: contig of 869 bp in length

\* 18042 18141: gap of 100 bp

\* 18142 19009: contig of 868 bp in length

\* 19010 19109: gap of 100 bp

\* 19110 19966: contig of 857 bp in length

\* 19967 20066: gap of 100 bp

\* 20067 20921: contig of 855 bp in length

\* 20922 21021: gap of 100 bp

\* 21022 21865: contig of 844 bp in length

\* 21866 21965: gap of 100 bp

\* 21966 22832: contig of 867 bp in length

\* 22833 22932: gap of 100 bp

\* 22933 23780: contig of 848 bp in length

\* 23781 23880: gap of 100 bp

\* 23881 24733: contig of 853 bp in length

\* 24734 24833: gap of 100 bp

\* 24834 25670: contig of 837 bp in length

\* 25671 25770: gap of 100 bp

\* 25771 26621: contig of 851 bp in length

\* 26622 26721: gap of 100 bp

\* 26722 27576: contig of 855 bp in length

\* 27577 27676: gap of 100 bp

\* 27677 28532: contig of 856 bp in length

\* 28533 28632: gap of 100 bp

\* 28633 29492: contig of 860 bp in length

\* 29493 29592: gap of 100 bp

\* 29593 30455: contig of 863 bp in length

\* 30456 30555: gap of 100 bp

\* 30556 31410: contig of 855 bp in length

\* 31411 31510: gap of 100 bp

\* 31511 32368: contig of 858 bp in length

\* 32369 32468: gap of 100 bp

\* 32469 33312: contig of 844 bp in length

\* 33313 33412: gap of 100 bp

\* 33413 34268: contig of 856 bp in length

\* 34269 34368: gap of 100 bp

\* 34369 35204: contig of 836 bp in length

\* 35205 35304: gap of 100 bp

\* 35305 36156: contig of 852 bp in length

\* 36157 36256: gap of 100 bp

\* 36257 37128: contig of 872 bp in length

\* 37129 37228: gap of 100 bp

\* 37229 38083: contig of 855 bp in length

\* 38084 38183: gap of 100 bp

\* 38184 39031: contig of 848 bp in length

\* 39032 39131: gap of 100 bp

\* 39132 40006: contig of 875 bp in length

\* 40007 40106: gap of 100 bp

\* 40107 40967: contig of 861 bp in length

\* 40968 41067: gap of 100 bp

\* 41068 41913: contig of 846 bp in length

\* 41914 42013: gap of 100 bp

\* 42014 42824: contig of 811 bp in length

\* 42825 42924: gap of 100 bp

\* 42925 43776: contig of 852 bp in length

\* 43777 43876: gap of 100 bp

\* 43877 44752: contig of 876 bp in length

\* 44753 44852: gap of 100 bp

\* 44853 45724: contig of 872 bp in length

\* 45725 45824: gap of 100 bp

\* 45825 46643: contig of 819 bp in length

\* 46644 46743: gap of 100 bp

\* 46744 47599: contig of 856 bp in length

\* 47600 47699: gap of 100 bp

\* 47700 48551: contig of 852 bp in length

\* 48552 48651: gap of 100 bp

\* 48652 49485: contig of 834 bp in length

\* 49486 49585: gap of 100 bp

\* 49586 50440: contig of 855 bp in length

\* 50441 50540: gap of 100 bp

\* 50541 51404: contig of 864 bp in length

\* 51405 51504: gap of 100 bp

\* 51505 52372: contig of 868 bp in length

\* 52373 52472: gap of 100 bp

\* 52473 53328: contig of 856 bp in length

\* 53329 53428: gap of 100 bp

\* 53429 54268: contig of 840 bp in length

\* 54269 54368: gap of 100 bp

\* 54369 55229: contig of 861 bp in length

\* 55230 55329: gap of 100 bp

Db	41705	CCCGGTTCTGTGTTTCCTCCCTTCFCGGAAGGCGTTGTCAAGGGGTAGGAGAAAGAGACGC	41764
Qy	3094	aaacacaaagtgaacacaggttaagaggtctccagtgacttacttggtggcggttatgtt	3153
Db	41765	ANACACAAAAGTGGAACAGGTAAAGAGCTCTCCAGTGACTTACTTGGCGGTATTGTT	41824
Qy	3154	ttgtttcaggccaagagaggtcttcgggaagtgcgtggtttcgggactttgatccggagc	3213
Db	41825	TTGTTTCGAAGCCCAAGGAG--CTTCGGGAGTGCTCGGTTCTTCGGGGACTTNTGATCCGACG	41882
Qy	3214	cc 3215	
Db	41883	cc 41884	
RESULT	13		
ACOL12230/LOCUS		ACOL12230	175064 bp DNA linear HTG 22-APR-2000
DEFINITION		Homo sapiens clone RP11-IM10, WORKING DRAFT SEQUENCE, 39 unordered pieces.	
ACCESSION		ACOL12230.3	GI-7637254
VERSION		ACOL12230.3	GI-7637254
KEYWORDS		HTG; HTGS_PHASE1; HTGS_DRAFT.	
SOURCE		human.	
ORGANISM		Homo sapiens	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS		1 (bases 1 to 175064)	
TITLE		Birren,B., Linton,L., Nusbaum,C. and Lander,E.	
JOURNAL		Homo sapiens, clone RP11-IM10	
REFERENCE		Unpublished	
AUTHORS		2 (bases 1 to 175064)	
		Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M., Baldwin,J., Barna,N., Beckerly,R., Boguslavskiy,L., Bouckgalter,B., Brown,A., Castie,A., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeAerillano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M., Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D., Galagan,J., Gardyna,S., Grant,G., Hagos,B., Hearford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Lehoczy,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X., Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.	
TITLE		Direct Submission	
JOURNAL		Submitted (31-OCT-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	
COMMENT		On Apr 22, 2000 this sequence version replaced gi:6454033.	
		All repeats were identified using RepeatMasker:	
		Smit, A.F.A. & Green, P. (1996-1997)	
		http://ftp.genome.washington.edu/RM/RepeatMasker.html	
		----- Genome Center	
		Center: Whitehead Institute/ MIT Center for Genome Research	
		Center code: WIBR	
		Web site: http://www-seq.wi.mit.edu	
		Contact: sequence_submissions@genome.wi.mit.edu	
		----- Project Information	
		Center project name: L2510	
		Center clone name: L_M10	
		----- Summary Statistics	
		Sequencing vector: M13; M77815; 100% of reads	
		Chemistry: Dye-terminator Big Dye; 100% of reads	
		Assembly program: Phrap; version 0.960731	
		Consensus quality: 117571 bases at least Q40	
		Consensus quality: 145749 bases at least Q30	
		Consensus quality: 160940 bases at least Q20	
		Insert size: 185000; agarose-fp	
		Insert size: 171264; sum-of-contigs	
		Quality coverage: 2.9 in Q20 bases; agarose-fp	
		Quality coverage: 3.2 in Q20 bases; sum-of-contigs	









```

Db 67162 GCGATAGTGTGTTTCGAGGCGCAAGAGGCTTCGAAAGTCTCGGTTTCGGGACTT 67103
QY 3203 tgatccgagagccacatcccc 3224
Db 67102 TGATCCGAGGCGCCACATCCCC 67081

RESULT 15
AC021345
LOCUS Homo sapiens clone RP11-24J9, LOW-PASS SEQUENCE SAMPLING.
DEFINITION AC021345
ACCESSION AC021345
VERSION AC021345.2 GI:9130845
KEYWORDS HTG; HTGS_PHASE0.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 90698)
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE Homo sapiens, clone RP11-24J9
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 90698)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Beda, F.,
Boguslavskiy, L., Boukhgalter, B., Brown, A., Burkett, G., Castle, A.,
Chapel, Y., Colangelo, M., Collins, S., Collamore, A., Cooke, P.,
DeAtellano, K., Dewar, K., Domino, M., Doyle, M., Fenestor, J.,
Ferrel, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J.,
Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
Landers, T., Lehoczy, J., Levine, R., Liu, C., Liu, G., Locke, K.,
Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,
McPheeters, R., Meldrum, J., Meneus, L., Morrow, J., Naylor, J.,
Norman, C. H., O'Connor, T., O'Donnell, P., Olivari, T. M., Peterson, K.,
Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,
Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J.,
Zimmer, A. and Zody, M.
Direct Submission
Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced gi:6705761.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L4483
Center clone name: 24_J_9
-----
* NOTE: This record contains 92 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
*
* 1 910: contig of 910 bp in length
* 911 1010: gap of 100 bp
* 1011 1073: contig of 863 bp in length
* 1874 1973: gap of 100 bp
* 1974 2824: contig of 851 bp in length

```

```

2825 2924: gap of 100 bp
2925 3802: contig of 878 bp in length
3803 3902: gap of 100 bp
3903 4816: contig of 914 bp in length
4817 4916: gap of 100 bp
4917 5759: contig of 843 bp in length
5760 5859: gap of 100 bp
5860 6764: contig of 905 bp in length
6765 6864: gap of 100 bp
6865 7747: contig of 883 bp in length
7748 7847: gap of 100 bp
7848 8755: contig of 908 bp in length
8756 8855: gap of 100 bp
8856 9753: contig of 898 bp in length
9754 9853: gap of 100 bp
9854 10757: contig of 904 bp in length
10758 10857: gap of 100 bp
10858 11732: contig of 875 bp in length
11733 11832: gap of 100 bp
11833 12739: contig of 907 bp in length
12740 12839: gap of 100 bp
12840 13710: contig of 871 bp in length
13711 13810: gap of 100 bp
13811 14684: contig of 874 bp in length
14685 14784: gap of 100 bp
14785 15662: contig of 878 bp in length
15663 15762: gap of 100 bp
15763 16677: contig of 915 bp in length
16678 16777: gap of 100 bp
16778 17678: contig of 901 bp in length
17679 17778: gap of 100 bp
17779 18679: contig of 901 bp in length
18680 18779: gap of 100 bp
18780 19632: contig of 853 bp in length
19633 19732: gap of 100 bp
19733 20634: contig of 902 bp in length
20635 20734: gap of 100 bp
20735 21620: contig of 886 bp in length
21621 21720: gap of 100 bp
21721 22579: contig of 859 bp in length
22580 22679: gap of 100 bp
22680 23568: contig of 889 bp in length
23569 23668: gap of 100 bp
23669 24554: contig of 886 bp in length
24555 24654: gap of 100 bp
24655 25521: contig of 867 bp in length
25522 25621: gap of 100 bp
25622 26487: contig of 866 bp in length
26488 26587: gap of 100 bp
26588 27464: contig of 877 bp in length
27465 27564: gap of 100 bp
27565 28466: contig of 902 bp in length
28467 28566: gap of 100 bp
28567 29464: contig of 898 bp in length
29465 29564: gap of 100 bp
29565 30447: contig of 883 bp in length
30448 30547: gap of 100 bp
30548 31453: contig of 906 bp in length
31454 31553: gap of 100 bp
31554 32452: contig of 899 bp in length
32453 32552: gap of 100 bp
32553 33447: contig of 895 bp in length
33448 33547: gap of 100 bp
33548 34435: contig of 888 bp in length
34436 34535: gap of 100 bp
34536 35433: contig of 898 bp in length
35434 35533: gap of 100 bp
35534 36440: contig of 907 bp in length
36441 36540: gap of 100 bp
36541 37423: contig of 882 bp in length
37423 37522: gap of 100 bp
37523 38402: contig of 880 bp in length
38403 38502: gap of 100 bp

```







GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 00:12:11 ; Search time 520.94 Seconds  
(without alignments)  
10648.732 Million cell updates/sec

Title: US-09-846-456-1  
Perfect score: 3231  
Sequence: 1 acaggcatgtgcagggtg.....gccccacatccccaccactt 3231

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_032802.\*

- 1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1980.DAT.\*
- 2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1981.DAT.\*
- 3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1982.DAT.\*
- 4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1983.DAT.\*
- 5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1984.DAT.\*
- 6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1985.DAT.\*
- 7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1986.DAT.\*
- 8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1987.DAT.\*
- 9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1988.DAT.\*
- 10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1989.DAT.\*
- 11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1990.DAT.\*
- 12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1991.DAT.\*
- 13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1992.DAT.\*
- 14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1993.DAT.\*
- 15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1994.DAT.\*
- 16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1995.DAT.\*
- 17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1996.DAT.\*
- 18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1997.DAT.\*
- 19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1998.DAT.\*
- 20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA1999.DAT.\*
- 21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2000.DAT.\*
- 22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT.\*
- 23: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2001B.DAT.\*
- 24: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	ID	Description
1	3127.8	96.8	183999	22 AAF92831 Human ABC1 genomic
2	1485.4	46.0	1643	22 AAF24681 Nucleotide sequenc
3	1485.4	46.0	1643	22 AAF24703 Nucleotide sequenc
4	336	10.4	763	22 AAF04729 Human cDNA clone (
5	336	10.4	1750	22 AAF17451 Human cDNA sequenc
6	330.8	10.2	20645	22 AAF05355 Human reproductive
7	330.6	10.2	5076	22 AAF16699 Human nervous syst
8	325.8	10.1	5075	22 AAF16701 Human nervous syst
9	323.8	10.0	36785	22 AAF82208 Human immune/haema

c 10	320.4	9.9	23934	22	ABAL9145	Human nervous syst
c 11	320.4	9.9	23934	22	AAAL36171	Human musculoskele
c 12	320.4	9.9	23934	22	AAAL36179	Human musculoskele
c 13	320.4	9.9	23934	22	AAAL04522	Human reproductive
c 14	320.4	9.9	23934	22	AAK28343	Genomic sequence #
c 15	320.4	9.9	23934	22	AAK71442	Human immune/haema
c 16	313.6	9.7	26390	22	AAK65971	Human immune/haema
c 17	312.8	9.7	9745	22	AAK36759	Human cardiovascular
c 18	312.8	9.7	12149	22	AAK36758	Human cardiovascular
c 19	312.4	9.7	8663	22	AAK30133	Human lung antigen
c 20	312.2	9.7	15275	22	AAK35975	Human cardiovascular
c 21	311.8	9.7	9731	22	AAK72933	Human immune/haema
c 22	311.8	9.7	9731	22	AAK85096	Human immune/haema
c 23	311.8	9.7	9733	22	AAK72935	Human immune/haema
c 24	311.8	9.7	9733	22	AAK85097	Human immune/haema
c 25	311.6	9.6	16555	22	AAK70102	Human immune/haema
c 26	311.6	9.6	16555	22	AAK73172	Human immune/haema
c 27	310.4	9.6	5797	22	ABAL5723	Human nervous syst
c 28	310.4	9.6	5797	22	AAAL03405	Human reproductive
c 29	309.8	9.6	23885	22	AAK70103	Human immune/haema
c 30	309.8	9.6	23885	22	AAK73173	Human immune/haema
c 31	309.8	9.6	66933	22	ABAB2625	Human HBM gene reg
c 32	309.8	9.6	72049	22	ABAB2623	Human HBM gene reg
c 33	307.2	9.5	15555	22	AAK73539	Human immune/haema
c 34	307.2	9.5	15558	22	AAK73538	Human immune/haema
c 35	306.2	9.5	23241	22	AAK97870	Human neuroblastom
c 36	306.2	9.5	23241	22	AAK97871	Human neuroblastom
c 37	304.8	9.4	15554	22	AAK73537	Human immune/haema
c 38	304.6	9.4	2743	22	AAK86888	Human immune/haema
c 39	304.2	9.4	8948	22	AAK67209	Human immune/haema
c 40	304	9.4	45000	22	AAAL2437	DNA encoding 1-ami
c 41	303.8	9.4	32191	22	AAK30497	DNA encoding novel
c 42	303.8	9.4	32191	22	AAAL06277	Human reproductive
c 43	303.8	9.4	46107	22	AAK71730	Human immune/haema
c 44	303.4	9.4	32224	22	AAK89986	Human digestive sy
c 45	303.2	9.4	26372	22	AAK77103	Human immune/haema

# ALIGNMENTS

## RESULT 1

AAAF92831  
ID AAF92831 standard; DNA; 183999 BP.

XX AAF92831;

AC AAF92831;

DT 17-MAY-2001 (first entry)

DE Human ABC1 genomic DNA.

DE Human ABC1 genomic DNA.

XX High density lipoprotein-cholesterol; HDL-C; cardiovascular; ABC1; ds.

XX Homo sapiens.

XX OS

XX WO200115676-A2.

XX PD 08-MAR-2001.

XX PF 01-SEP-2000; 2000WO-IB01492.

XX PR 01-SEP-1999; 99US-0151977.

XX PR 15-MAR-2000; 2000US-0526193.

XX PR 23-JUN-2000; 2000US-0213958.

XX (UYBR-) UNIV BRITISH COLUMBIA.

XX (XENO-) XENON GENETICS INC.

PI Hayden MR, Brooks-Wilson AR, Pimstone SN, Clee SM;

XX WPI; 2001-244356/25.

PT Treating a lower than normal high density lipoprotein-cholesterol



Db 27680 gtttgacgtgagttttgcccagaataaagtgacatttagttgttgcttgatgac 27739  
QY 1846 ttaaatatttagac--atggtgttagccctgcattctactctctgttctttttttgccc 1903  
Db 27740 ttaaatatttagacataatggtgttagccctgcattctactctctgttctttttttgccc 27799  
QY 1904 cctccagtggtttggttagttttgtctccctacagcccaaaagcagagaagtgtgag 1963  
Db 27800 cctccagtggtttggttagttttgtctccctacagcccaaaagcagagaagtgtgag 27859  
QY 1964 gcttgagtggtctacataattttacagagactgcaattctctggtgcaattcacaaatgt 2023  
Db 27860 gcttgagtggtctacataattttacagagactgcaattctctggtgcaattcacaaatgt 27919  
QY 2024 atacaactaaatacaagctgctgtttttatcacagggagctgataataataatgaaa 2083  
Db 27920 atacaactaaatacaagctgctgtttttatcacagggagctgataataataatgaaa 27979  
QY 2084 ttaaaaggggctgctacataattgtctgtgtttttgtttgtttgtttgtttgtttt 2143  
Db 27980 ttaaaaggggctgctacataattgtctgtgtttttgtttgtttgtttgttttntnn 28039  
QY 2144 tttttgtttttgtgctctctctctctcaatttatgaagaagcagtaagatttctc 2203  
Db 28040 tnnntgtttttgtgctctctctctctcaatttatgaagaagcagtaagatttctc 28099  
QY 2204 tcgggtctctgagggacctgggagctcagctcagctgggaatctccaagcagtagtgccc 2263  
Db 28100 tcgggtctctgagggacctgggagctcagctcagctgggaatctccaagcagtagtgccc 28159  
QY 2264 tatcaaaatcaaatccagtggtttgtgggggaaacaaaagcagccattaccacaggg 2323  
Db 28160 tatcaaaatcaaatccagtggtttgtgggggaaacaaaagcagccattaccacaggg 28219  
QY 2324 actgtccgcttccctcaaccagctagcccttgaaaggaacaaaagcagacagcaaa 2383  
Db 28220 actgtccgcttccctcaaccagctagcccttgaaaggaacaaaagcagacagcaaa 28279  
QY 2384 aatgattggctcctgagggagattcagctagagctctctccccaatccctccctc 2443  
Db 28280 aatgattggctcctgagggagattcagctagagctctctccccaatccctccctc 28339  
QY 2444 cggctgaggaactaacaaggaagaaaaaaatttcggaaagcagagatttagaagaagca 2503  
Db 28340 cggctgaggaactaacaaggaagaaaaaaatttcggaaagcagagatttagaagaagca 28399  
QY 2504 attccactgtgctccttgctgcccgggaacgtggactagagactgctgcccagccc 2563  
Db 28400 attccactgtgctccttgctgcccgggaacgtggactagagactgctgcccagccc 28459  
QY 2564 agcccagcgtctccgcgcgtcttaggcgcgcggccggcgaggggagggagcgacaga 2623  
Db 28460 agcccagcgtctccgcgcgtcttaggcgcgcggccggcgaggggagggagcgacaga 28519  
QY 2624 ccgcggaacctaaagacacctgtgtaccctccaccccccccccacccacccacccctccc 2683  
Db 28520 ccgcggaacctaaagacacctgtgtaccctccaccccccccccacccacccacccctccc 28579  
QY 2684 ccaactcctagatgtgtgtggtggcggtgaactgcgcctgttaagggggcgcccccgg 2743  
Db 28580 ccaactcctagatgtgtgtggtggcggtgaactgcgcctgttaagggggcgcccccgg 28639  
QY 2744 ctcccagtgctctctgctgagtgactgaactacataaagagccggagggggcgcg 2803  
Db 28640 ctcccagtgctctctgctgagtgactgaactacataaagagccggagggggcgcg 28699  
QY 2804 gaggaggagagacagagctttgaccgtagtagtaacctctgcgtcgtgtagcgcgaatct 2863  
Db 28700 gaggaggagagacagagctttgaccgtagtagtaacctctgcgtcgtgtagcgcgaatct 28759  
QY 2864 ataaaggaactagtcctccgggcaaaaaccccgtaattgcgagcagagtgagtgggccgg 2923  
Db 28760 ataaaggaactagtcctccgggcaaaaaccccgtaattgcgagcagagtgagtgggccgg 28819

QY 2924 gacccgcagagccgagccgaccccttctcccggtgctgcggcagggcagggcgaggagct 2983  
Db 28820 gacccgcagagccgagccgaccccttctcccggtgctgcggcagggcagggcgaggagct 28879  
QY 2984 ccgcgcacacacagagccggtctcagggcggtttgctcctgttttttcccggttctg 3043  
Db 28880 ccgcgcacacacagagccggtctcagggcggtttgctcctgttttttcccggttctg 28939  
QY 3044 ttttctcccttctccggaaggtgtcgaagggttaggagaaagagacgacacacaaaa 3103  
Db 28940 ttttctcccttctccggaaggtgtcgaagggttaggagaaagagacgacacacaaaa 28999  
QY 3104 gtgaaacacaggttaagaggtctccagtgacttacttggcggttattgtttgttcgag 3163  
Db 29000 gtgaaacacaggttaagaggtctccagtgacttacttggcggttattgtttgttcgag 29059  
QY 3164 gcaagagagcttcgggaagtgcgtcgttttcggggactttgatccggagccacacatccc 3223  
Db 29060 gcaagagagcttcgggaagtgcgtcgttttcggggactttgatccggagccacacatccc 29119  
QY 3224 caccactt 3231  
Db 29120 caccactt 29127

## RESULT 2

AAF24681

ID AAF24681 standard; DNA; 1643 BP.

XX AAF24681;

XX 20-APR-2001 (first entry)

XX Nucleotide sequence of the 5' flanking region of the human ABC1 gene.

Human; adenosine triphosphate binding cassette protein 1; ABC1;  
XW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;  
KW chromosome 9q22-9q31; heart disease; hypercholesterolemia;  
KW atherosclerosis; cholesterol transport; ss.

XX Homo sapiens.

XX WO200078972-A2.

XX 28-DEC-2000.

XX 16-JUN-2000; 2000WO-US16765.

XX 18-JUN-1999; 99US-0140264.

XX 14-SEP-1999; 99US-0153872.

XX 19-NOV-1999; 99US-0166573.

XX (CVTH-) CV THERAPEUTICS INC.

XX Lawn RM, Wade D, Garvin M;

XX WPI; 2001-137812/14.

Adenosine triphosphate (ATP) binding cassette (ABC) polynucleotide,  
useful for the development of agents for the treatment of heart disease  
and other disorders associated with hypercholesterolemia and  
atherosclerosis -

PS Claim 1; Page 143-144; 215pp; English.

XX The present sequence represents the 5' flanking region of the human  
CC adenosine triphosphate (ATP) binding cassette protein (ABC) 1 gene. ABC1  
CC resides in cell membranes and utilizes ATP hydrolysis to transport a wide  
CC variety of substrates across the plasma membrane. ABC1 is a pivotal  
CC protein in the apolipoprotein-mediated mobilisation of intracellular  
CC cholesterol stores. ABC1 is defective in Tangier disease, a genetic  
CC disorder characterised by abnormal HDL-cholesterol metabolism. The ABC1





PN WO200078971-A2.  
 XX 28-DEC-2000.  
 XX 16-JUN-2000; 2000WO-US16591.  
 XX 18-JUN-1999; 99US-0140264.  
 PR 14-SEP-1999; 99US-0153872.  
 PR 19-NOV-1999; 99US-0166573.  
 XX (CVTH-) CV THERAPEUTICS INC.  
 PA (UNIW ) UNIV WASHINGTON.  
 XX Lawn RM, Wade D, Oram JF, Garvin M;  
 PI WPI; 2001-137811/14.  
 DR Adenosine triphosphate (ATP) binding cassette protein (ABC) 1  
 PT polynucleotides and polypeptides, useful for treatment of heart disease  
 PT and other disorders associated with hypercholesterolemia and  
 PT atherosclerosis -  
 XX  
 PS Disclosure; Page 138-139; 211pp; English.  
 XX  
 CC The present sequence represents the 5' flanking region of the human  
 CC adenosine triphosphate (ATP) binding cassette protein (ABC) 1 gene. ABC1  
 CC resides in cell membranes and utilizes ATP hydrolysis to transport a wide  
 CC variety of substrates across the plasma membrane. ABC1 is a pivotal  
 CC protein in the apolipoprotein-mediated mobilisation of intracellular  
 CC cholesterol stores. ABC1 is defective in Tangier disease, a genetic  
 CC disorder characterised by abnormal HDL-cholesterol metabolism. The ABC1  
 CC gene is localised to chromosome 9q22-9q31. The ABC1 genes and proteins  
 CC are useful for developing pharmaceutical agents for the treatment of  
 CC heart disease and other disorders associated with hypercholesterolemia  
 CC and atherosclerosis. The genes are useful for developing screening assays  
 CC to screen for compounds that regulate the expression of genes associated  
 CC with cholesterol transport. The genes and proteins are also useful for  
 CC are also useful as diagnostic indicators of cardiovascular disease and  
 CC other disorders associated with hypercholesterolemia.  
 XX  
 SQ Sequence 1643 BP; 370 A; 413 C; 457 G; 403 T; 0 other;

Query Match 46.0%; Score 1485.4; DB 22; Length 1643;  
 Best Local Similarity 98.1%; Pred. No. 5e-305;  
 Matches 1615; Conservative 0; Mismatches 6; Indels 26; Gaps 10;

QY 1345 gctccacatgcactccaggcgtgtggcctctctctatgcgtctgtctgagtgtt 1404  
 DB 16 ggtccacatgcactccaggcgtgtggcctctctctatgcgtctgtctgagtgtt 73

QY 1405 gatagaaccactgatgtgagctacgtggcgttgagccgtggcctggagatcctgttgactg 1464  
 DB 74 gatagaaccactgatgtgagctacgtggcgttgag-cgtggcctggagatcctgttgactg 132

QY 1465 tagcatgagggcgtctgtcagctgaatgctcagtcaggtggtggagttcttgaat 1524  
 DB 133 tagcatgagggcgtctgtcagctgaatgctcagtcaggtggtggagttcttgaat 191

QY 1525 atgatggagctggaggtgggaagagaagtagccttgaggcagctctctcatgcacatca 1584  
 DB 192 atgatggagctggaggtgggaagagaagtagccttgaggcagctctctcatgcacatca 251

QY 1585 ttctggccaaactcaggctcaaacgtggaagagcttaaaatggaatggccttcaagggt 1644  
 DB 252 ttctggccaaactcaggctcaaacgtggaagagcttaaaatggaatggccttcaagggt 311

QY 1645 ggctacaaaggtatctttgtcaaggtgagagacctgtggcctccacgtgcacttccagg 1704  
 DB 312 ggctacaaaggtatctttgtcaaggtgagagacctgtggcctccacgtgcacttccagg 371

QY 1705 gctcgtgtggcctctctacgtgtgtcgtcgtgagttcttatgaat---ccttcagggc 1761  
 DB 1437 taaacagagcgggaacggggcgaggagggagagacagagcgtttgaccgatagtaa 1496

DB 372 gctcgtgtggcctctctacgtgtgtcgtcgtgagttcttatgaatcctccttcagggc 430  
 QY 1762 agattcaatattagactcttcacagtttgacctgaagtttggccagaataagtgacatt 1821  
 DB 431 agattcaatattagactcttcacagtttggacctgagtttggccagaataagtgacatt 490  
 QY 1822 tagtttgtggtgcttgagtgagcttaaaattattagac--atggtgtgtaggctgcatt 1879  
 DB 491 tagtttgtggtgcttgagtgagcttaaaattattagacatattggtgtaggctgcatt 550  
 QY 1880 cctactcttgccttttttttttgcctccagtggttttgggtagttttgtc-ccccacag 1938  
 DB 551 cctactcttgccttttttttttgcctccagtggttttgggtagttttgtcctccctacag 610  
 QY 1939 ccaaaaggcaaacagagaagttggaggtctgagtggtggtacataattttacacagctgcaa 1998  
 DB 611 ccaaaaggcaaacagagaagttggaggtctgagtggtggtacataattttacacagctgcaa 670  
 QY 1999 ttctctggtgctgacttcacaaatgtatacaaaactaaatacaagtcctgtttttatcac 2058  
 DB 671 ttctctggtgctgacttcacaaatgtatacaaaactaaatacaagtcctgtttttatcac 730  
 QY 2059 agggaggtgatcaataataataaaagggtggtggt-ccatattgttctgtgttt 2117  
 DB 731 agggaggtgatcaataataataaaagggtggtggtggtcctcatattgttctgtgttt 790  
 QY 2118 ttgtttgtttgtttgttt 2177  
 DB 791 ttg-----ttgtgttt 841  
 QY 2178 tgaagagaacagtaagatttctctcggttctctgagggacctggggagctcaggct 2237  
 DB 842 tgaagagaacagtaagatttctctcggttctctgagggacctggggagctcaggct 901  
 QY 2238 gggaattctcaaggcagtagtgctctcaaaaatacaaaagtcaggtttgtgggggaa 2297  
 DB 902 gggaattctcaaggcagtagtgctctcaaaaatacaaaagtcaggtttgtgggggaa 961  
 QY 2298 acaaaagagcaccattaccagagagactgtccgcttccctcaccctcagcagcctcaggcct 2357  
 DB 962 acaaaagcagccattaccagagagactgtccgcttccctcaccctcagcagcctcaggcct 1021  
 QY 2358 ttgaagagaacaaaaagacaagcaaaatgattggcgtctgagggagattcagctcaga 2417  
 DB 1022 ttgaagagaacaaaaagacaagcaaaatgattggcgtctgagggagattcagctcaga 1081  
 QY 2418 gctctctctccccaatccctcctcctcggtgaggaactaacaaggaagaaaaaaattg 2477  
 DB 1082 gctctctctccccaatccctcctcctcggtgaggaactaacaaggaagaaaaaaattg 1141  
 QY 2478 cggaaagcaggaatttagaggaagcaaaattccactggtgcttggctgctggcggaacgtgg 2537  
 DB 1142 cggaaagcaggaatttagaggaagcaaaattccactggtgcttggctgctggcggaacgtgg 1201  
 QY 2538 actagagagctcgtcgcgcagcccgagcctcagcagcttcccgcgcttcttagccgcggcg 2597  
 DB 1202 actagagagctcgtcgcgcagcccgagcctcagcagcttcccgcgcttcttagccgcggcg 1261  
 QY 2598 gcccgggcggggaaagggaacgcagaccgagccctcaagacacctgctgtaacctccac 2657  
 DB 1262 gcccgggcggggaaagggaacgcagaccgagccctcaagacacctgctgtaacctccac 1321  
 QY 2658 ccccaaccccaaccccaacccctcccccaactcccccaactccctagatgtgtctggtggcggtgaaag 2717  
 DB 1322 -----ccccaccccaacccctcccccaactcccccaactccctagatgtgtctggtggcggtgaaag 1376  
 QY 2718 tcgccgttttaaggcgcccgcccgctccactgctcactgcttctctgctgagtgactgaactaca 2777  
 DB 1377 tcgccgttttaaggcgcccgcccgctccactgctcactgcttctctgctgagtgactgaactaca 1436  
 QY 2778 taaacagagcggcggaagggggcgaggagggagagacagagcgtttgaccgatagtaa 2837  
 DB 1437 taaacagagcggcggaagggggcgaggagggagagacagagcgtttgaccgatagtaa 1496









PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250391.  
PR 01-DEC-2000; 2000US-0251160.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-541565/60.  
XX  
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,  
XX useful for preventing, diagnosing and/or treating nervous system  
XX cancers and metastases -  
XX  
XX Disclosure; SEQ ID NO 9030; 1701pp + Sequence Listing; English.  
XX  
XX The invention relates to novel genes (ABA11004-ABA21534) and proteins  
XX (AB14678-AB18001) useful for preventing, treating or ameliorating  
XX medical conditions e.g. by protein or gene therapy. The genes are  
XX isolated from a range of human tissues disclosed in the specification.  
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful  
XX in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
XX and ovarian cancer and other cancers of the adrenal gland, bone, bone  
XX marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
XX (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
XX haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
XX disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
XX colitis; (c) cardiovascular disorders such as myocardial ischaemias;  
XX (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
XX epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
XX and parasitic infections.  
XX Note: the sequence data for this patent did not form part of the  
XX printed specification, but was obtained in electronic format directly  
XX from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 5076 BP; 1435 A; 1052 C; 1191 G; 1398 T; 0 other;  
XX  
XX Query Match 10.2%; Score 330.6; DB 22; Length 5076;  
XX Best Local Similarity 65.0%; Pred. No. 2.1e-60;  
XX Matches 587; Conservative 0; Mismatches 299; Indels 17; Gaps 6;  
XX  
XX QY 375 ggccgggacaaatggctgaatccccagcacttggaggcctgagccgagagga 434

QY 08-NOV-2000; 2000US-0246613.  
QY 17-NOV-2000; 2000US-0249207.  
QY 17-NOV-2000; 2000US-0249208.  
QY 17-NOV-2000; 2000US-0249209.  
QY 17-NOV-2000; 2000US-0249210.  
QY 17-NOV-2000; 2000US-0249211.  
QY 17-NOV-2000; 2000US-0249212.  
QY 17-NOV-2000; 2000US-0249213.  
QY 17-NOV-2000; 2000US-0249214.  
QY 17-NOV-2000; 2000US-0249215.  
QY 17-NOV-2000; 2000US-0249216.  
QY 17-NOV-2000; 2000US-0249217.  
QY 17-NOV-2000; 2000US-0249218.  
QY 17-NOV-2000; 2000US-0249244.  
QY 17-NOV-2000; 2000US-0249245.  
QY 17-NOV-2000; 2000US-0249264.  
QY 17-NOV-2000; 2000US-0249265.  
QY 17-NOV-2000; 2000US-0249297.  
QY 17-NOV-2000; 2000US-0249299.  
QY 17-NOV-2000; 2000US-0249300.  
QY 01-DEC-2000; 2000US-0250391.  
QY 01-DEC-2000; 2000US-0251160.  
QY 05-DEC-2000; 2000US-0251030.  
QY 05-DEC-2000; 2000US-0251988.  
QY 05-DEC-2000; 2000US-0256719.  
QY 06-DEC-2000; 2000US-0251479.  
QY 08-DEC-2000; 2000US-0251856.  
QY 08-DEC-2000; 2000US-0251868.  
QY 08-DEC-2000; 2000US-0251869.  
QY 08-DEC-2000; 2000US-0251989.  
QY 08-DEC-2000; 2000US-0251990.  
QY 11-DEC-2000; 2000US-0254097.  
QY 05-JAN-2001; 2001US-0259678.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-541565/60.  
XX  
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,  
XX useful for preventing, diagnosing and/or treating nervous system  
XX cancers and metastases -  
XX  
XX Disclosure; SEQ ID NO 9030; 1701pp + Sequence Listing; English.  
XX  
XX The invention relates to novel genes (ABA11004-ABA21534) and proteins  
XX (AB14678-AB18001) useful for preventing, treating or ameliorating  
XX medical conditions e.g. by protein or gene therapy. The genes are  
XX isolated from a range of human tissues disclosed in the specification.  
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful  
XX in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
XX and ovarian cancer and other cancers of the adrenal gland, bone, bone  
XX marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
XX (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
XX haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
XX disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
XX colitis; (c) cardiovascular disorders such as myocardial ischaemias;  
XX (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
XX epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
XX and parasitic infections.  
XX Note: the sequence data for this patent did not form part of the  
XX printed specification, but was obtained in electronic format directly  
XX from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 5076 BP; 1435 A; 1052 C; 1191 G; 1398 T; 0 other;  
XX  
XX Query Match 10.2%; Score 330.6; DB 22; Length 5076;  
XX Best Local Similarity 65.0%; Pred. No. 2.1e-60;  
XX Matches 587; Conservative 0; Mismatches 299; Indels 17; Gaps 6;  
XX  
XX QY 375 ggccgggacaaatggctgaatccccagcacttggaggcctgagccgagagga 434

Db 3222 GGCCTGGTGCATTGGCTCATGCTGTAAATCCAGCACTTTGGAGGCTGAGCGCGGA 3163  
QY 435 tcgcttgagctccaagatttgagaccagcctggataacatggcacaacccctgctctaca 494  
Db 3162 TCACCTGAGGTCAGGAGTTTGAGACAAGCTTCCCAACATGGTGAACCC-CCATCTCTTA 3104  
QY 495 aaaaaatacaaaaattagatgggtgtgggtggcatgcaacctgggtcccagctacttggg 554  
Db 3103 CTAAATAATCAAAAATAGCCAGGCATGGTGGCAGCGCTGTAAATCCAGCTACTCAGG 3044  
QY 555 aggttaaggtgggagatcgcttgagccagggagtgcaagtctacactagagccatattg 614  
Db 3043 AGGCTGAGGCAGAGAATTCCTTGAACCCAGAGGTGGAGGTCCAGTGCAGCCGAGATCA 2984  
QY 615 gatactgcaactccagcctg---ggtagacagagcagaccctgtctcaaaaaaagaaa 671  
Db 2983 CGCCACTGCAATCCAGCCTGGATAGATAGAGCAGGACTCTGTCTCAAAAAAAGAGAG 2924  
QY 672 tgaagagagaaaagagagagagagagatgaggggagagagagaggggggggag 731  
Db 2923 AGAGAGAAAAAAGAGCTTTGATAAGGAACCTTGGCATAACCAATAATTTGGAAGATAA 2864  
QY 732 gaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaagga 791  
Db 2863 AAGTCAAGGAAGCTTTCAGGATCTTTTACCAATATTTTAAAGTCTAAAGTGACAACCTA 2804  
QY 792 ggcagaaagactttacgtataatgtctcatcatgtgtgtgtcaagtttgaccccaaaccc 851  
Db 2803 GCTAGCCAAATGAGAAATGTGTTAATGGACTTATCAGGCTTAAGACT--GCCTTAAGCTG 2747  
QY 852 aattattgaccaggttattctttgactgaggaagggggggtccgcctcctctgggcttg 911  
Db 2746 CTGGGTAGAATACAGGGGCACCTTGTCTTCCCGTGGAGGCTCTTCAGGTGAGCTTCTTATCA 2687  
QY 912 ggccttagaagctcatctctggtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 971  
Db 2686 ACTTCTGACCAACCTGGCCCAATGCTTCTTGGAAATTAATTTTTTTTTTTTTTTTTTTT 2627  
QY 972 gacacgagctctgtctgtcactcaggtggagtgagtgagtgagtgagtgagtgagtgagtg 1031  
Db 2626 GAGATGGAGTCTGCTCTGTCCACCCAGGCTGGAATGCAACGGTACCATCTCAGTCTCAGTG 2567  
QY 1032 taacctgtgctcccggttcaagcgattctcctgctcagcctcctga-----gat 1083  
Db 2566 CAACCTCTGCTCCCGGTTCAAGCGATTCTCTGCTCAGCTCCCTGAGTAGTAGTGGGAT 2507  
QY 1084 aacagcgcccgccaccacacatctggctaaattttgtatttttagtaagactgggttca 1143  
Db 2506 TACAGGATGTGCCACCCAGCCCGGCTAAATTTTGTGA-TTTTAGTAGAGATGGGTTTCG 2448  
QY 1144 tcatgttggccaggttgggttcgaactcctgacctgagctgagctgagctgagctgagctg 1203  
Db 2447 CCATGTTGGCCAGGCTGATCTCGAACTCTCGACCTCAGCTGAGTGTGCTGCTGCTTGGCCTC 2388  
QY 1204 ccaag-tgctggattacaggaatgagcactgcccagctcagatccatcccttct 1262  
Db 2387 CCAAAGTGGAGGATTTACAGGCAATGAGCCAGCCGCTGCCAGACAAATTTTTTTTTTTTTT 2328  
QY 1263 aag 1265  
Db 2327 AAG 2325

RESULT 8  
ABAI6701/C  
ID ABAI6701 standard; DNA; 5075 BP.  
XX  
XX ABAI6701;  
XX  
XX 23-JAN-2002 (first entry)  
XX  
XX Human nervous system related polynucleotide SEQ ID NO 9032.

XX Human; nootropic; neuroprotective; cytotstatic; dermatological; virucide;  
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;  
KW antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;  
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;  
KW antiallergic; antidiabetic; antilulcer; anticonvulsant; antifungal;  
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.  
XX Homo sapiens.  
OS  
XX WO200159063-A2.  
XX  
XX 16-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US01334.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220863.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 23-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0228927.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0234984.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0242221.  
PR 08-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-024474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 01-DEC-2000; 2000US-0250390.  
PR 01-DEC-2000; 2000US-0250391.  
PR 01-DEC-2000; 2000US-0251160.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.

PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-541565/60.

XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,  
 PT useful for preventing, diagnosing and/or treating nervous system  
 PT cancers and metastases -

XX Disclosure; SEQ ID NO 9032; 1701pp + Sequence Listing; English.

XX The invention relates to novel genes (AB11004-ABA21534) and proteins  
 CC (AB114678-AB18001) useful for preventing, treating or ameliorating  
 CC medical conditions e.g. by protein or gene therapy. The genes are  
 CC isolated from a range of human tissues disclosed in the specification.  
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful  
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone  
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
 CC haemolytic anemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;  
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
 CC and parasitic infections.

CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 5075 BP; 1437 A; 1054 C; 1187 G; 1397 T; 0 other;

Query Match 10.1%; Score 325.8; DB 22; Length 5075;

Best Local Similarity 64.7%; Pred. No. 2.2e-59;

Matches 584; Conservative 0; Mismatches 302; Indels 17; Gaps 6;

QY 375 gcccgggcaaatgctgaatgctgaatccagcactttggaggctgagccagaga 434

Db 3221 GGCTGGTGGCTATGCTGCTGTAATCCAGCACTTTGGAGGCTGAGCGGCGGA 3162

QY 435 tcgcttgagctccagagtttgagaccagcctggataacatggcaaacccctgtcttaca 494

Db 3161 TCACCTGAGGTGAGGATTTGAGACAAAGCTAGCCAAACATGGTGAACCC-CCATCTCTA 3103

QY 495 aaaaaatacaaaaataagatgggtgtgtgtgcatgcacactgtgtccagctacttggg 554

Db 3102 CTAATAATACAAAATAGCCAGGCAATGTGGCAGCGCTGTAAATCCAGCTACTCAGG 3043

QY 555 agcctaagtgaggagatcgcttgagccagcaggagagcaactgatactgagcagatgag 614

Db 3042 AGGCTGAGGCAAGAAGATTCCTTGAAACCCAGGAGGTGGAGGTGGCAGTGGCGGAGATCA 2983

QY 615 gatcactgactccagcctg---ggtagacagagaagaccctgtctcaaaaaaagaaa 671

Db 2982 CGCCATGCTATCCAGCCTGGATGATAGATAGAGCGAGACTGTGTCTCAAAAAAGAGAG 2923

QY 672 tgaagagagaagaagaagagagagagagagatgaggggagagggagggggag 731

Db 2922 AGAGAGAAAAAAGACTCTTTGATAGGAACTTGGCATACTAAATAATTTGGAAGATAAA 2863

QY 732 gaaggaggaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 791

Db 2862 AAGTCAGGAGAGTTCAGATCTTTTACCAATATTTTAAAGCTAAAGTCAACACTA 2803

QY 792 ggcagaaagactttacgttaattgtctcatgtgtgtgtcaagtgttgcaagtttgaccccaaaccc 851  
 Db 2802 CTAAGCCAAATGAGAAATGTGATTTAATGGACTTATCAGGCTAAGACTG---CCTTAAGCT 2746  
 QY 852 aattattgaccaaggttatctttgactgaggcaagggggttcgctctctctgggcccctg 911  
 Db 2745 GCTGGTAGAATCAGGGGCACTTCTCTCCGTTGGAGGGTCTTCAGGTGACGTTCTTATC 2686  
 QY 912 ggcattagaagatcattctctggcctttctgagatccatccctttctttttttttt 971  
 Db 2685 AACTTCTGACCAACTGGCCAAATGCTTCTCTGGAATAATTTTTTTTTTTTTTTTTT 2626  
 QY 972 gacacgaggtctgtctgtcactcaggtgagtgagtgagtgagtgagtgagtgagtgagtgag 1031  
 Db 2625 GAGATGGAGTCTGCTGTCTCACCAGGCTGGAATGCAACGGTACCATCTCAGCTCACTG 2566  
 QY 1032 taactctgctcccggttcaagcattctctctgctcagcctcctga-----gat 1083  
 Db 2565 CAACCTCTGCTCCCGGTTTCAAGGGATTCTCTCCCTCAGCCCTCCTAGTAGTGGGAT 2506  
 QY 1084 aacaggggccgcaccacatctgctaattttgtatttttttttttttttttttttttttttt 1143  
 Db 2505 TACAGGCAATGTCACCAACCCGCGCTAAATTTTGTGTA-TTTAGTAGAGATGGGTTTCG 2447  
 QY 1144 tcatgttgccaggtgtgtgttctgaaactcctcagcctgagtgagtgagtgagtgagtgagtgag 1203  
 Db 2446 CCATGTTGGCCAGGCTGATCTCGAACTCCTGACCTCAGGTGATCTGCTGCTTGGCCTC 2387  
 QY 1204 ccaag-tgctggattacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1262  
 Db 2386 CCAAAGTTGCAAGGATTAACAGGCATGAGCCACCGTGGCCAGACAAATTTTTTTTTTTTTT 2327  
 QY 1263 aag 1265  
 Db 2326 AAG 2324

RESULT 9

AAK82208

ID AAK82208 standard; DNA; 36785 BP.

XX AAK82208;

AC AAK82208;  
 XX 07-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:37020.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
 KW cytostatic; gene therapy; vaccine; metastasis; ds.

XX Homo sapiens.

XX WO200157182-A2.

XX 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US01354.

XX 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184684.

XX 02-MAR-2000; 2000US-0186350.

XX 16-MAR-2000; 2000US-0189874.

XX 17-MAR-2000; 2000US-0190076.

XX 18-APR-2000; 2000US-0198123.

XX 19-MAY-2000; 2000US-0205515.

XX 07-JUN-2000; 2000US-0209467.

XX 28-JUN-2000; 2000US-0214886.

XX 30-JUN-2000; 2000US-0215135.

XX 07-JUL-2000; 2000US-0216647.

XX 07-JUL-2000; 2000US-0216880.

XX 11-JUL-2000; 2000US-0217487.



PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226686.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234233.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.

PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 01-DEC-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 03-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483426/52.

Nucleic acids encoding human immune/hematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis -

Disclosure; SEQ ID NO 37020; 3071pp + Sequence Listing; English.

AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I) amino acid sequences given in AAK62170 to AAK61921. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins, and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to supplement the patient's own production of (I). Additionally, (I) polynucleotides may be used to produce the secreted (I), by inserting the nucleic acids into a host cell and culturing the cell to express the

CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169  
CC represent sequences used in the exemplification of the present invention.  
XX  
SQ Sequence 36785 BP; 9938 A; 8156 C; 8241 G; 10450 T; 0 other;

Query Match 10.0%; Score 323.8; DB 22; Length 36785;  
Best Local Similarity 63.5%; Pred. No. 9.8e-59;  
Matches 566; Conservative 0; Mismatches 307; Indels 18; Gaps 4;  
Qy 370 aacagcgccggcacaatgctgaatgctgaatccacagcaactttggagcgtgagccca 429  
Dy 17728 aaaaagccagagtggtgctcatgcttaattctagcaactttggagcgtgagccgg 17787  
Qy 430 gaggatcgttgagctccagagtttgagaccagcctggataacatgccaacccctgtct 489  
Dy 17788 gtggatcacctgagatcagagtttgagataagcctggccaacatggtgaaacccgtgt 17847  
Qy 490 ctacaaaaatacaaaaattagatggtgtgtggtgcatgcaactgtgttccagctac 549  
Dy 17848 ctactaaaaata---caaaattagcaggtgtgtggtgcatgcaactgtgttccagctac 17904  
Qy 550 ttggagggctaaagtgaggagctgcttgagccagggagtgcaactgacactgagccat 609  
Dy 17905 ttccgagggctgagcagagaaatcagttgaaacccggaggtgaggttgagtgagctga 17964  
Qy 610 gattggtatcactccagcctgggtgtagacagagcagaccctgtctcaaaaaaaga 659  
Dy 17965 gattgacacttgactccagcctgggcaacagagcgaaactgtgtctcaaaagaaaa 18024  
Qy 670 aatcaagaagaagaagaagagagagagagagagagagagagagagagagagagag 724  
Dy 18025 gaaaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 18084  
Qy 725 gggggaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaag 784  
Dy 18085 ttacttagtgaccttggtggcaaaactttgacccaaattctcaactatataagaagag 18144  
Qy 785 aacagagggcagaagactttagtaaatgctcaatgctgaatgctgaatgctgaatg 844  
Dy 18145 attagattagttttagctggtttatcaattgttttctcagaatttttaggggtttta 18204  
Qy 845 aaaaaccaattattgacaaagttatttttctgacagagcagaggggtccgtctctcg 904  
Dy 18205 gaaatgcttgaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaagga 18264  
Qy 905 ggccttgggtttagaagctcatctctgcttctgagatccatccctttctttttat 964  
Dy 18265 tgcctaaccaagaatagttccacttttaccagttttattactgatactctttttttt 18324  
Qy 965 tttcttgacacagagctctgctc--tgcactcaaggttgagtgagtgagtgatctc 1022  
Dy 18325 tttttttgacagaggttctgctcttcttcttcttcttcttcttcttcttcttctc 18384  
Qy 1023 gactcaactgaactctgctccgggttcaacgattctcctgctcagctc----- 1076  
Dy 18385 tccctatggaactctgctccaggttccaggttccaggttccaggttccaggttccaggt 18444  
Qy 1077 --ctgagataacagggcccgccaccacatctgctaaatttttttttttttttagtaagac 1134  
Dy 18445 agctggattacagggccctgcccacacgcccggcgaatttttttttttttttttagagac 18504  
Qy 1135 tgggtttcactaatgttgccaggtgtgttctgcaactcctgacctgaggtgagtgccac 1194  
Dy 18505 ggggtttcactaatgttgccaggtgtgttctgcaactcctgacctgaggtgagtgccac 18564  
Qy 1195 ctggcctcccaagtgctggattacagcagcagcagcagcagcagcagcagcagcagc 1245  
Dy 18565 ctacgctcccaagtgctggattacagcagcagcagcagcagcagcagcagcagcagc 18615

RESULT 10  
ID ABA19145/C  
XX ABA19145 standard; DNA; 23934 BP.  
AC ABA19145;  
XX  
DT 23-JAN-2002 (first entry)  
XX  
DE Human nervous system related polynucleotide SEQ ID NO 11476.  
XX  
KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;  
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnery;  
KW antiparinsonian; antisclerol; antianemic; antiarthritic; cancer;  
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;  
KW antiallergic; antidiabetic; antilucer; anticonvulsant; antifungal;  
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200159063-A2.  
XX  
PD 16-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01334.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226688.  
PR 23-AUG-2000; 2000US-0227182.  
PR 30-AUG-2000; 2000US-0227009.  
PR 01-SEP-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.

PR	08-SEP-2000;	2000US-0231243.	PR	17-NOV-2000;	2000US-0249264.
PR	08-SEP-2000;	2000US-0231244.	PR	17-NOV-2000;	2000US-0249265.
PR	08-SEP-2000;	2000US-0231413.	PR	17-NOV-2000;	2000US-0249297.
PR	08-SEP-2000;	2000US-0231414.	PR	17-NOV-2000;	2000US-0249299.
PR	08-SEP-2000;	2000US-0232080.	PR	17-NOV-2000;	2000US-0249300.
PR	08-SEP-2000;	2000US-0232081.	PR	01-DEC-2000;	2000US-0250391.
PR	12-SEP-2000;	2000US-0231968.	PR	01-DEC-2000;	2000US-0251160.
PR	14-SEP-2000;	2000US-0232397.	PR	05-DEC-2000;	2000US-0251030.
PR	14-SEP-2000;	2000US-0232398.	PR	05-DEC-2000;	2000US-0251988.
PR	14-SEP-2000;	2000US-0232399.	PR	05-DEC-2000;	2000US-0256719.
PR	14-SEP-2000;	2000US-0232400.	PR	06-DEC-2000;	2000US-0251479.
PR	14-SEP-2000;	2000US-0232401.	PR	08-DEC-2000;	2000US-0251856.
PR	14-SEP-2000;	2000US-0233063.	PR	08-DEC-2000;	2000US-0251868.
PR	14-SEP-2000;	2000US-0233064.	PR	08-DEC-2000;	2000US-0251869.
PR	14-SEP-2000;	2000US-0233065.	PR	08-DEC-2000;	2000US-0251989.
PR	21-SEP-2000;	2000US-0234223.	PR	08-DEC-2000;	2000US-0251990.
PR	21-SEP-2000;	2000US-0234274.	PR	11-DEC-2000;	2000US-0254097.
PR	23-SEP-2000;	2000US-0234997.	PR	05-JAN-2001;	2001US-0259678.
PR	25-SEP-2000;	2000US-0234998.	XX		
PR	26-SEP-2000;	2000US-0234984.	PA	(HUMA-) HUMAN GENOME SCI INC.	
PR	27-SEP-2000;	2000US-0235484.	XX		
PR	27-SEP-2000;	2000US-0235834.	PI	Rosen CA, Barash SC, Ruben SM;	
PR	27-SEP-2000;	2000US-0235836.	XX		
PR	29-SEP-2000;	2000US-0236327.	DR	WPI; 2001-541565/60.	
PR	29-SEP-2000;	2000US-0236367.	XX		
PR	29-SEP-2000;	2000US-0236368.	XX		
PR	29-SEP-2000;	2000US-0236369.	PT	Nucleic acids encoding 3224 human nervous system antigen polypeptides,	
PR	29-SEP-2000;	2000US-0236370.	PT	useful for preventing, diagnosing and/or treating nervous system	
PR	02-OCT-2000;	2000US-0236802.	PT	cancers and metastases -	
PR	02-OCT-2000;	2000US-0237037.	XX		
PR	02-OCT-2000;	2000US-0237038.	PS	Disclosure; SEQ ID NO 11476; 1701pp + Sequence Listing; English.	
PR	02-OCT-2000;	2000US-0237039.	XX		
PR	02-OCT-2000;	2000US-0237040.	XX		
PR	13-OCT-2000;	2000US-0239935.	CC	The invention relates to novel genes (ABAI1004-ABA21534) and proteins	
PR	13-OCT-2000;	2000US-0239937.	CC	(ABBI4678-ABBI8001) useful for preventing, treating or ameliorating	
PR	20-OCT-2000;	2000US-0240960.	CC	medical conditions e.g. by protein or gene therapy. The genes are	
PR	20-OCT-2000;	2000US-0241785.	CC	isolated from a range of human tissues disclosed in the specification.	
PR	20-OCT-2000;	2000US-0241786.	CC	The nucleic acids, proteins, antibodies and (ant)agonists are useful	
PR	20-OCT-2000;	2000US-0241787.	CC	in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast	
PR	20-OCT-2000;	2000US-0241808.	CC	and ovarian cancer and other cancers of the adrenal gland, bone, bone	
PR	20-OCT-2000;	2000US-0241809.	CC	marrow, breast, gastrointestinal tract, liver, lung, or urogenital;	
PR	20-OCT-2000;	2000US-0241826.	CC	(b) immune disorders e.g. Addison's disease, allergies, autoimmune	
PR	20-OCT-2000;	2000US-0242221.	CC	haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's	
PR	01-NOV-2000;	2000US-0244617.	CC	disease, multiple sclerosis, rheumatoid arthritis and ulcerative	
PR	08-NOV-2000;	2000US-0246474.	CC	colitis; (c) cardiovascular disorders such as myocardial ischaemia;	
PR	08-NOV-2000;	2000US-0246475.	CC	(d) wound healing; (e) neurological diseases e.g. cerebral anoxia and	
PR	08-NOV-2000;	2000US-0246476.	CC	epilepsy; and (f) infectious diseases such as viral, bacterial, fungal	
PR	08-NOV-2000;	2000US-0246477.	CC	and parasitic infections.	
PR	08-NOV-2000;	2000US-0246478.	CC	Note: The sequence data for this patent did not form part of the	
PR	08-NOV-2000;	2000US-0246523.	CC	printed specification, but was obtained in electronic format directly	
PR	08-NOV-2000;	2000US-0246524.	CC	from WIPO at ftp.wipo.int/pub/published_pct_sequences.	
PR	08-NOV-2000;	2000US-0246525.	XX		
PR	08-NOV-2000;	2000US-0246526.	SQ	Sequence 23934 BP; 7785 A; 4761 C; 4365 G; 7023 T; 0 other;	
PR	08-NOV-2000;	2000US-0246527.			
PR	08-NOV-2000;	2000US-0246528.			
PR	08-NOV-2000;	2000US-0246532.			
PR	08-NOV-2000;	2000US-0246509.			
PR	08-NOV-2000;	2000US-0			





```
Db 3190 tagggactgagagatctg-----gaagcataaagaaggaatgagacagagtacactctg 3244
Qy 991 tcaactcagctgagtgagtgatctcgactcactgtaactctgctcccggt 1050
Db 3245 tcaccagctgagtgagtgatctcgactcactgtaactcaccctcggt 3304
Qy 1051 tcaagcattctcctcagctc-----ctgagataaacagcgcccgccaccac 1102
Db 3305 tcaagcattctcagctcagctcccaagtagctgggtattacagcgcatgccacgat 3364
Qy 1103 atctgctaatcttatttttagtaagactgggtttcatctgttgccaggttgg 1162
Db 3365 gctcgctaatcttatttttagtagagacaggttttaccatgttgccaggtgat 3424
Qy 1163 ttgaactcctgactgaggtgagctgcccacttggcctcccaagtgtggattaca 1222
Db 3425 ctcaattctgactcattgtatccacccactcggcctcccaagtgtggattaca 3484
Qy 1223 ggcagcactgcccagc 1244
Db 3485 ggcagcactcaccacccgc 3506

RESULT 12
AAL36179/C
ID AAL36179 standard; DNA; 23934 BP.
XX
AC AAL36179;
XX
DT 08-JAN-2002 (first entry)
XX
DE Human musculoskeletal system related polynucleotide SEQ ID NO 2544.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein;
KW musculoskeletal system; ds.
XX
OS Homo sapiens.
XX
PN WO20015367-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01338.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
```



AAL04522  
 ID AAL04522 standard; DNA; 23934 BP.  
 XX  
 AC AAL04522;  
 DT 21-NOV-2001 (first entry)  
 XX  
 XX Human reproductive system related antigen DNA SEQ ID NO: 7210.  
 DE  
 XX Human; reproductive system related antigen; reproductive system disorder;  
 KW cancer; gene therapy; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200155320-A2.  
 PN  
 XX  
 PD 02-AUG-2001.  
 XX  
 PF 17-JAN-2001; 2001WO-US01339.  
 XX  
 PR 31-JAN-2000; 2000US-0179065.  
 PR 04-FEB-2000; 2000US-0180628.  
 PR 24-FEB-2000; 2000US-0184664.  
 PR 02-MAR-2000; 2000US-0186350.  
 PR 16-MAR-2000; 2000US-0189874.  
 PR 17-MAR-2000; 2000US-0190076.  
 PR 18-APR-2000; 2000US-0198123.  
 PR 19-MAY-2000; 2000US-0205515.  
 PR 07-JUN-2000; 2000US-0209467.  
 PR 28-JUN-2000; 2000US-0214886.  
 PR 30-JUN-2000; 2000US-0215135.  
 PR 07-JUL-2000; 2000US-0216647.  
 PR 07-JUL-2000; 2000US-0216880.  
 PR 11-JUL-2000; 2000US-0217487.  
 PR 11-JUL-2000; 2000US-0217496.  
 PR 14-JUL-2000; 2000US-0218290.  
 PR 26-JUL-2000; 2000US-0220963.  
 PR 26-JUL-2000; 2000US-0220964.  
 PR 14-AUG-2000; 2000US-0224518.  
 PR 14-AUG-2000; 2000US-0224519.  
 PR 14-AUG-2000; 2000US-0225213.  
 PR 14-AUG-2000; 2000US-0225214.  
 PR 14-AUG-2000; 2000US-0225266.  
 PR 14-AUG-2000; 2000US-0225267.  
 PR 14-AUG-2000; 2000US-0225268.  
 PR 14-AUG-2000; 2000US-0225270.  
 PR 14-AUG-2000; 2000US-0225447.  
 PR 14-AUG-2000; 2000US-0225757.  
 PR 14-AUG-2000; 2000US-0225758.  
 PR 14-AUG-2000; 2000US-0225759.  
 PR 18-AUG-2000; 2000US-0226279.  
 PR 22-AUG-2000; 2000US-0226681.  
 PR 22-AUG-2000; 2000US-0226868.  
 PR 22-AUG-2000; 2000US-0227182.  
 PR 23-AUG-2000; 2000US-0227009.  
 PR 30-AUG-2000; 2000US-0228924.  
 PR 01-SEP-2000; 2000US-0229287.  
 PR 01-SEP-2000; 2000US-0229343.  
 PR 01-SEP-2000; 2000US-0229344.  
 PR 01-SEP-2000; 2000US-0229345.  
 PR 05-SEP-2000; 2000US-0229509.  
 PR 05-SEP-2000; 2000US-0229513.  
 PR 06-SEP-2000; 2000US-0230437.  
 PR 06-SEP-2000; 2000US-0230438.  
 PR 08-SEP-2000; 2000US-0231242.  
 PR 08-SEP-2000; 2000US-0231243.  
 PR 08-SEP-2000; 2000US-0231244.  
 PR 08-SEP-2000; 2000US-0231413.  
 PR 08-SEP-2000; 2000US-0231414.  
 PR 08-SEP-2000; 2000US-0232080.  
 PR 08-SEP-2000; 2000US-0232081.  
 PR 12-SEP-2000; 2000US-0231968.  
 PR 14-SEP-2000; 2000US-0232397.  
 PR 14-SEP-2000; 2000US-0232398.  
 PR 14-SEP-2000; 2000US-0232399.  
 PR 14-SEP-2000; 2000US-0232400.  
 PR 14-SEP-2000; 2000US-0232401.  
 PR 14-SEP-2000; 2000US-0233063.  
 PR 14-SEP-2000; 2000US-0233064.  
 PR 14-SEP-2000; 2000US-0233065.  
 PR 21-SEP-2000; 2000US-0234223.  
 PR 21-SEP-2000; 2000US-0234224.  
 PR 21-SEP-2000; 2000US-0234274.  
 PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.





PR	14-AUG-2000;	2000US-0225267.	PR	08-NOV-2000;	2000US-0246526.	XX	PA	(HUMA-) HUMAN GENOME SCI INC.	XX	Sequence 23934 BP; 7785 A; 4761 C; 4365 G; 7023 T; 0 other;
PR	14-AUG-2000;	2000US-0225267.	PR	08-NOV-2000;	2000US-0246526.	XX	PI	Rosen CA, Barash SC, Ruben SM;	XX	
PR	14-AUG-2000;	2000US-0225268.	PR	08-NOV-2000;	2000US-0246527.	XX	PI	WPI; 2001-476224/51.	XX	
PR	14-AUG-2000;	2000US-0225270.	PR	08-NOV-2000;	2000US-0246528.	DR	XX		XX	
PR	14-AUG-2000;	2000US-0225447.	PR	08-NOV-2000;	2000US-0246532.	XX	XX	Isolated polypeptide for treating, preventing and/or prognosing	XX	
PR	14-AUG-2000;	2000US-0225757.	PR	08-NOV-2000;	2000US-0246609.	PT	PT	disorders related to the respiratory system including respiratory	PT	
PR	14-AUG-2000;	2000US-0225758.	PR	08-NOV-2000;	2000US-0246610.	XX	XX	cancers and also for testing and detection e.g. diagnosis -	XX	
PR	14-AUG-2000;	2000US-0225759.	PR	08-NOV-2000;	2000US-0246611.	PS	PS	Disclosure; SED ID No 777; 546pp; English.	PS	
PR	18-AUG-2000;	2000US-0228279.	PR	08-NOV-2000;	2000US-0246613.	XX	XX	The present invention relates to the isolation of novel human	XX	
PR	22-AUG-2000;	2000US-0228681.	PR	17-NOV-2000;	2000US-0249207.	CC	CC	respiratory antigens (AAU17685-AAU17975), and cDNA and genomic	CC	
PR	22-AUG-2000;	2000US-0228688.	PR	17-NOV-2000;	2000US-0249208.	CC	CC	sequences encoding for these polypeptides. The sequences of the	CC	
PR	22-AUG-2000;	2000US-0227182.	PR	17-NOV-2000;	2000US-0249209.	CC	CC	invention are useful for preventing, treating and/or prognosing	CC	
PR	23-AUG-2000;	2000US-0227009.	PR	17-NOV-2000;	2000US-0249210.	CC	CC	disorders related to the respiratory system including throat	CC	
PR	30-AUG-2000;	2000US-0228924.	PR	17-NOV-2000;	2000US-0249211.	CC	CC	disorders (e.g. vocal cord paralysis, tonsillitis, and laryngitis),	CC	
PR	01-SEP-2000;	2000US-0229287.	PR	17-NOV-2000;	2000US-0249212.	CC	CC	lung disorders e.g. pneumonia, allergic disorders e.g. asthma,	CC	
PR	01-SEP-2000;	2000US-0229343.	PR	17-NOV-2000;	2000US-0249213.	CC	CC	pleurisy, cystic fibrosis, emphysema, nose disorders and cancers of	CC	
PR	01-SEP-2000;	2000US-0229344.	PR	17-NOV-2000;	2000US-0249214.	CC	CC	the respiratory tissues e.g. lung cancer. The polynucleotide sequences	CC	
PR	01-SEP-2000;	2000US-0229345.	PR	17-NOV-2000;	2000US-0249215.	CC	CC	of the invention are useful in gene therapy and antisense therapy.	CC	
PR	05-SEP-2000;	2000US-0229509.	PR	17-NOV-2000;	2000US-0249216.	CC	CC	AA528161-AAS28764 represent genomic sequences encoding for novel	CC	
PR	05-SEP-2000;	2000US-0229513.	PR	17-NOV-2000;	2000US-0249217.	CC	CC	human respiratory antigens.	CC	
PR	06-SEP-2000;	2000US-0230437.	PR	17-NOV-2000;	2000US-0249218.	CC	CC	Note: The sequence data for this patent did not form part of the printed	CC	
PR	06-SEP-2000;	2000US-0230438.	PR	17-NOV-2000;	2000US-0249219.	CC	CC	specification, but was obtained in electronic format directly from WIPO	CC	
PR	08-SEP-2000;	2000US-0231242.	PR	17-NOV-2000;	2000US-0249220.	XX	XX	at ftp.wipo.int/pub/published_pct_sequences.	XX	
PR	08-SEP-2000;	2000US-0231243.	PR	17-NOV-2000;	2000US-0249221.	XX	XX	Sequence 23934 BP; 7785 A; 4761 C; 4365 G; 7023 T; 0 other;	XX	
PR	08-SEP-2000;	2000US-0231244.	PR	17-NOV-2000;	2000US-0249222.	XX	XX		XX	
PR	08-SEP-2000;	2000US-0231413.	PR	17-NOV-2000;	2000US-0249223.	XX	XX		XX	
PR	08-SEP-2000;	2000US-0231414.	PR	17-NOV-2000;	2000US-0249224.	XX	XX		XX	
PR	08-SEP-2000;	2000US-0232080.	PR	17-NOV-2000;	2000US-0249225.	XX	XX		XX	
PR	08-SEP-2000;	2000US-0232081.	PR	17-NOV-2000;	2000US-0249226.	XX	XX		XX	
PR	12-SEP-2000;	2000US-0231968.	PR	17-NOV-2000;	2000US-0249227.	XX	XX		XX	
PR	14-SEP-2000;	2000US-0232397.	PR	17-NOV-2000;	2000US-0249228.	XX	XX		XX	
PR	14-SEP-2000;	2000US-0232398.	PR	17-NOV-2000;	2000US-0249229.	XX	XX		XX	
PR	14-SEP-2000;	2000US-0232399.	PR	17-NOV-2000;	2000US-0249230.	XX	XX		XX	
PR	14-SEP-2000;	2								

Best Local Similarity 64.28; Pred. No. 4.6e-58;

Matches 592; Conservative 0; Mismatches 296; Indels 34; Gaps 6;

Qy 348 tctctccattttaagatgaagaacagcgccggcacaatgctgaatgcctgttaataccca 407  
 Db 21341 TTTCTCACAATAAGACCTTACTTAGGAGCGCAGACGCTGGTGCATATGCTTAAATATCCCA 21282  
 Qy 408 gcaactttggaggtgagcgagaggtcgcttgagctccagagcttgagacagcctgg 467  
 Db 21281 GCACCTTTGGAGCGCGAGGCTGGCAGATCGCTTGAGTCGAGGAGTTTGAGACAGCATGG 21222  
 Qy 468 ataacatggcaaacacctgtctctacaaaaaaatacaaaaaatagatgggtggtggc 527  
 Db 21221 GCAACATGGCAAAACCCCATCTCTACAAAAAATA---CAAAATTAACCAAGGTGTGTGGC 21165  
 Qy 528 atgcacctgtgtccagctacttgagctgaagctgaagtgagagctcttgagcccaagg 587  
 Db 21164 ACATACCTGTGTAGTCCAGCTACTGAGGAGGCTGAGGACACAGAAATGCTTGAGCCTGGGA 21105  
 Qy 588 agtcaagtctacatgagccatgattgagctgactccagcctgggtgagacagagca 647  
 Db 21104 GCGGAGGTTGACGTGAGCCGAGATCATGCCACTGCACCTTAGCCTGGG-CGACAGAGCA 21046  
 Qy 648 agacctctctcaaaaaaaagaataaagaagaagaagaagaagagagagagagaga 707  
 Db 21045 AGACTCTGTATCAAAAAAATAAAAAAATAAAAAAATAAACTTAGGAACCTCTAAAGTCAA 20986  
 Qy 708 tgaggag 767  
 Db 20985 GTAGAAGTGATCAGATATTTATCTGCTATAATAATAAAGGTAGAAAGTAAATCATATTT 20926  
 Qy 768 aagaaaaaaacaagatg-----aaacagagcgagcaagaagactttacgttaattgctcat 820  
 Db 20925 TAAGAAGGAACAATGGGTTTCATACCTCATGAAGAGTTAAATGAGTTCAAAAGAGGGA 20866  
 Qy 821 catgtggtgtgaagtttgaccccaaaacccaat-----ttattgacccaaggtta 870  
 Db 20865 CATAAAATTTTCTACTTTGGGAGCAGACAGAAAAATCATGGAATGTTCTCTGAAAAATTTT 20806  
 Qy 871 ttcttgactgagcaaggggtccgtctctctggtggtggtggtggtggtggtggtggt 930  
 Db 20805 ATTACAGATGAAGCAGAGAGAACTCTGGTTGAGGAGAGGCAATGGGAGCGCCACGG 20746  
 Qy 931 ctggccttctgagatccatcccttttcttttttttttttttttttttttttttttttt 990  
 Db 20745 TAGGACTGGAGAGATCTG-----GAAGCATAAAGAAAGATGAGACAGAGTATCACTCTG 20691  
 Qy 991 tcaactaaggtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1050  
 Db 20690 TCACCCAGGCTGGAGTGAGTAGCACAATCTTGGCTCACTGCAACCTCCACCTCCTGGGT 20631  
 Qy 1051 tcaagcgattctctgctcagctc-----ctgagataacagcgcccgccacccac 1102  
 Db 20630 TCAAGGATTCATGCTCAGCTCCAGTACTCCAGTACTGGGATTACAGCGGATGCCAGAT 20571  
 Qy 1103 atcggcctaatttttatt 1162  
 Db 20570 GCCTGGTAAATTTTGTATTTTGTATTTTGTATTTTGTATTTTGTATTTTGTATTTTGT 20511  
 Qy 1163 ttcaactctctgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1222  
 Db 20510 CTCCAATTCCTGAGCTCATGTGATTCACCCACCTCCGCGCTCCCAAGTGTGGGATTACA 20451  
 Qy 1223 ggcagtgagcactgcgcccagc 1244  
 Db 20450 GGCATGAGCCATCACACCGGC 20429

RESULT 15

AAK71442

ID AAK71442 standard; DNA; 23934 BP.

XX

AC AAK71442;

XX DT 06-NOV-2001 (first entry)  
 XX DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:26254.  
 XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
 XX KW cytostatic; gene therapy; vaccine; metastasis; ds.  
 OS Homo sapiens.  
 XX PN WO200157182-A2.  
 XX PD 09-AUG-2001.  
 XX PF 17-JAN-2001; 2001WO-US01354.  
 XX PR 31-JAN-2000; 2000US-0179065.  
 PR 24-FEB-2000; 2000US-0180628.  
 PR 24-FEB-2000; 2000US-0184664.  
 PR 02-MAR-2000; 2000US-0186350.  
 PR 16-MAR-2000; 2000US-0189874.  
 PR 17-MAR-2000; 2000US-0190076.  
 PR 18-APR-2000; 2000US-0198123.  
 PR 19-MAY-2000; 2000US-0205515.  
 PR 07-JUN-2000; 2000US-0209467.  
 PR 28-JUN-2000; 2000US-0214886.  
 PR 30-JUN-2000; 2000US-0215135.  
 PR 07-JUL-2000; 2000US-0216647.  
 PR 07-JUL-2000; 2000US-0216880.  
 PR 11-JUL-2000; 2000US-0217487.  
 PR 14-JUL-2000; 2000US-0217496.  
 PR 14-JUL-2000; 2000US-0218290.  
 PR 26-JUL-2000; 2000US-0220963.  
 PR 26-JUL-2000; 2000US-0220964.  
 PR 14-AUG-2000; 2000US-0224518.  
 PR 14-AUG-2000; 2000US-0224519.  
 PR 14-AUG-2000; 2000US-0225213.  
 PR 14-AUG-2000; 2000US-0225214.  
 PR 14-AUG-2000; 2000US-0225266.  
 PR 14-AUG-2000; 2000US-0225267.  
 PR 14-AUG-2000; 2000US-0225268.  
 PR 14-AUG-2000; 2000US-0225270.  
 PR 14-AUG-2000; 2000US-0225447.  
 PR 14-AUG-2000; 2000US-0225757.  
 PR 14-AUG-2000; 2000US-0225758.  
 PR 14-AUG-2000; 2000US-0225759.  
 PR 18-AUG-2000; 2000US-0228279.  
 PR 22-AUG-2000; 2000US-0228681.  
 PR 22-AUG-2000; 2000US-0228688.  
 PR 23-AUG-2000; 2000US-0227182.  
 PR 30-AUG-2000; 2000US-0227009.  
 PR 01-SEP-2000; 2000US-0229287.  
 PR 01-SEP-2000; 2000US-0229343.  
 PR 01-SEP-2000; 2000US-0229344.  
 PR 05-SEP-2000; 2000US-0229345.  
 PR 05-SEP-2000; 2000US-0229509.  
 PR 06-SEP-2000; 2000US-0229513.  
 PR 06-SEP-2000; 2000US-0230437.  
 PR 06-SEP-2000; 2000US-0230438.  
 PR 08-SEP-2000; 2000US-0231242.  
 PR 08-SEP-2000; 2000US-0231243.  
 PR 08-SEP-2000; 2000US-0231244.  
 PR 08-SEP-2000; 2000US-0231413.  
 PR 08-SEP-2000; 2000US-0231414.  
 PR 08-SEP-2000; 2000US-0232080.  
 PR 08-SEP-2000; 2000US-0232081.  
 PR 12-SEP-2000; 2000US-0231968.  
 PR 14-SEP-2000; 2000US-0232397.  
 PR 14-SEP-2000; 2000US-0232398.  
 PR 14-SEP-2000; 2000US-0232399.  
 PR 14-SEP-2000; 2000US-0232400.  
 PR 14-SEP-2000; 2000US-0232401.

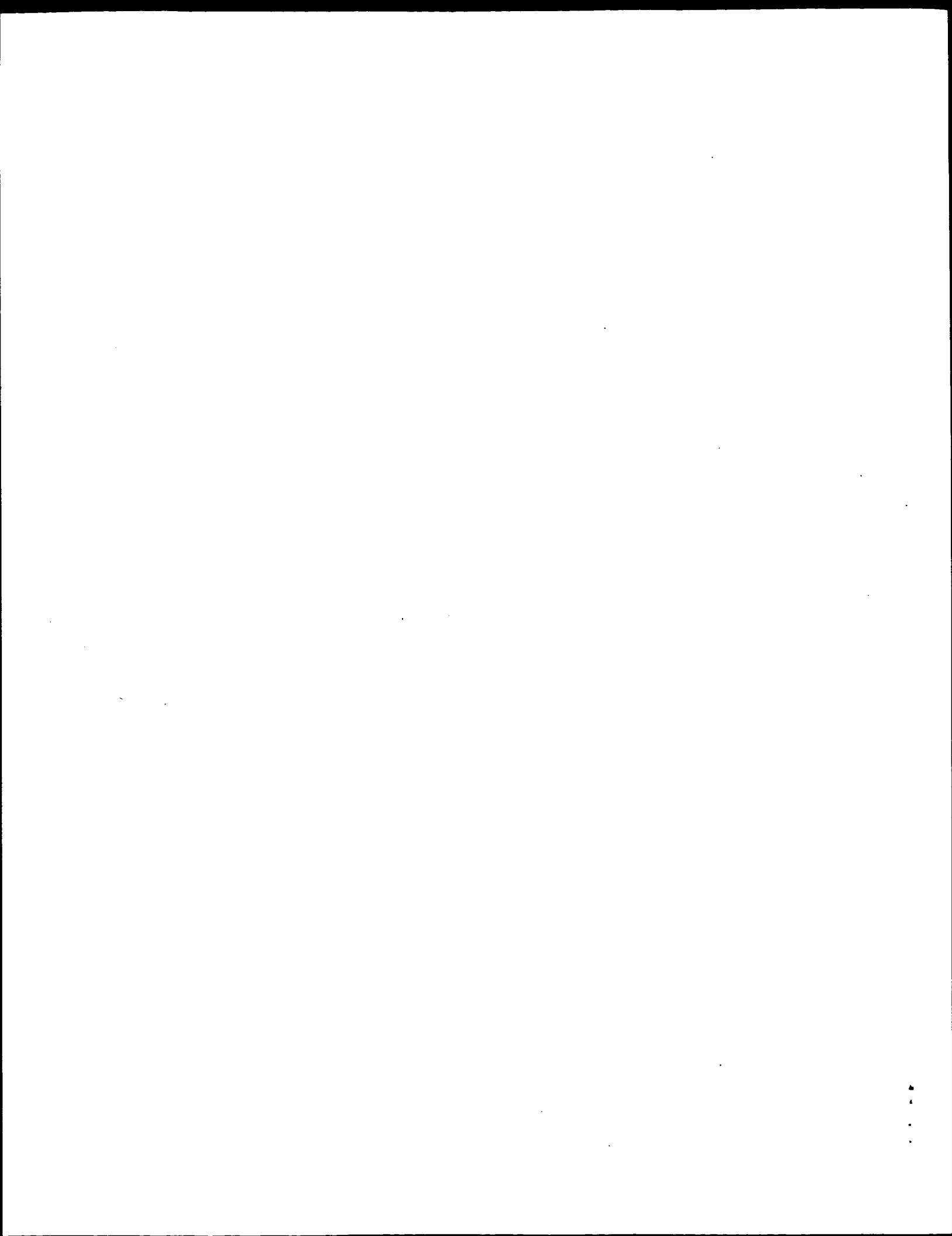
[illegible]

```

Db 3010 taagaaggaacaattggggtttcaactcatgaagattaaatgagttcaaaagagggga 3069
QY 821 catgtggtttgaagtttgaccccaaaacccaat-----ttattgaccaagggtta 870
Db 3070 cataaaattttctactttgggagacagaaaaaatcatggaatgttctctgaaaaatttt 3129
QY 871 ttctttgactgaggcaaggggtccgtctctctggtggttggccttttagaaaagctcatct 930
Db 3130 atcagagatgaagccagaagaactctgggttgaggagaaggcattggggagccagcgg 3189
QY 931 ctggcctttctgagatccatcccttcttctttttttcttgacacgagctctgctctg 990
Db 3190 tagggactggagagatcg----gaagcataagaagaaatgagacagagatcactctg 3244
QY 991 tcaactcaggtcgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1050
Db 3245 tcaccaggtcgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 3304
QY 1051 tcaagcattctcctgctcagcctc-----ctgagataacagagcgccgccaccac 1102
Db 3305 tcaagcattctcagctcagcctcagcctcagcctcagcctcagcctcagcctcagc 3364
QY 1103 atctggctaattttgtatttttagtaaaagactgggtttctcatctgttgccaggttgg 1162
Db 3365 gctggctaattttgtatttttagtagagacaggggttttaccatgttgccaggtgat 3424
QY 1163 ttcgaaactcctgacctgaggtgagctgcccacaccttgccctcccaaaagtgtggattaca 1222
Db 3425 ctccaatcctgacctcagctgagctgcccacaccttgccctcccaaaagtgtggattaca 3484
QY 1223 ggcattgagccactgcgccacc 1244
Db 3485 ggcattgagccactcacaccggc 3506

```

Search completed: September 20, 2002, 03:08:22  
Job time: 10571 sec









NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 15297  
TYPE: DNA  
ORGANISM: Human  
US-09-817-180-3

Query Match 8.4%; Score 271.8; DB 4; Length 15297;  
Best Local Similarity 62.9%; Pred. No. 1.8e-57;  
Matches 577; Conservative 0; Mismatches 287; Indels 53; Gaps 8;

QY 376 gccgggcaacaatggcctaagctgtaataccagcactctggagagcctgagcagagagat 435  
Db 14390 gctgggtgcagtgctcacacataatccagcactctggagagcctgagcagagat 14449

QY 436 cgctgagctccagagttgagaccagcctggataacatggcaaacctctgtctac-a 494  
Db 14450 cactgaggtcagagctcagaccagctggcccaacatagtaaaactcgtctgtcta 14509

QY 495 aaaaaatacaaaaattagatgggtggtggcagcagcctggaacacagcctgacttggg 554  
Db 14510 aatacaaaaataatagcagcgtggtggtgacacgtgaatccacagcactctgg 14569

QY 555 agcctaaggtggagagctcgttgagccagagagagagagagagagagagagagagat 614  
Db 14570 agcctgagaaggag 14629

QY 615 gatcactgactccagcctggtgagacagag-----caagacctgtctcaaaaaaaa 667  
Db 14630 cgccattgactccagcctggtgagacagagagagagagagagagagagagagagag 14689

QY 668 gaaatgaaagagaaagaaagagagagagagagagagagagagagagagagagagag 727  
Db 14690 aaaaacagcaaaaataatctatttgaagagagatgagagagagagagagagagag 14749

QY 728 ggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaag 787  
Db 14750 aaaaaggaaggt 14809

QY 788 caggggagaaagacttaagtaaatgctcatatgctgtgtgtgtgtgtgtgtgtgtgtgt 847  
Db 14810 -----cctcaatggaaatggtagaaagcaga 14836

QY 848 acccaattattgaccaggttattcttgaactgaggaagagagagagagagagagagagag 907  
Db 14837 actgaaaaactctgctagg-taggaatggttaggtgtgtgtgtgtgtgtgtgtgtgt 14895

QY 908 ctgggctttgaaagctcactctgccccttctgagatccatcccttcttttttttttt 967  
Db 14896 caacgcctgagggaaaaaacagcctaagatgaaatgtcttttttttttttttttttt 14954

QY 968 tctgacagcagctctgtctgtctgactcagcctggagtgagtgagtgagtgagtgagtgag 1027  
Db 14955 ttttgagatgagctgcgctgtgtgacagcctggagtgagtgagtgagtgagtgagtgag 15014

QY 1028 actgtaacctgtcctccgggttcaagcagatctcctgctcagcctcctga----- 1080  
Db 15015 actgtaacctgtcctccgggttcaagcagatctcctgctcagcctcctgagtagctg 15074

QY 1081 -gataacagcgcgcgcac-cacatctggctaatatttttttttttttttttttttttttt 1138  
Db 15075 ggaatacagggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 15134

QY 1139 ttctcatctgtggcagaggt 1197  
Db 15135 ttctcaactgtgtgagcgt 15194

QY 1198 ggcctcccaagctgtggattacaggaatgagcagcagcagcagcagcagcagcagcagc 1257  
Db 15195 ggcctcccaagctgtggattacaggaatgagcagcagcagcagcagcagcagcagcagc 15254

QY 1258 ttctaaaggggcaaacag 1274  
Db 15255 ttctaaagggaatgaag 15271

RESULT 4  
US-09-341-587-7  
; Sequence 7, Application US/09341587  
; Patent No. 6346606  
; GENERAL INFORMATION:  
; APPLICANT: Mollenhauer, Jan  
; TITLE OF INVENTION: Protein Containing an SRCR Domain  
; FILE REFERENCE: 4121-108  
; CURRENT APPLICATION NUMBER: US/09/341.587  
; CURRENT FILING DATE: 1999-08-31  
; EARLIER APPLICATION NUMBER: PCT/DE98/00096  
; EARLIER FILING DATE: 1998-01-09  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 28720  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-341-587-7

Query Match 8.3%; Score 269.4; DB 4; Length 28720;  
Best Local Similarity 61.5%; Pred. No. 9.7e-57;  
Matches 589; Conservative 0; Mismatches 296; Indels 72; Gaps 7;

QY 355 attttaaagatgaagaaacagccgggcacaaatggcctaagctgctgaatccccagcacttt 414  
Db 18740 atcttataaacaacacactggccaggatgctgctcaggtgtgtaataccccagcacttt 18799

QY 415 gggagagctgagccagagagctgctgagctccagagtttgagaccagcctgagataacat 474  
Db 18800 gggag 18859

QY 475 gcaaaacccctgctctaca-aaaaaaatacaaaaaattagatgggtgtgtgtgtgtgtgtgt 533  
Db 18860 agggagagccctctctctacaaaaagtaaaaaattagctgggtgtgtgtgtgtgtgtgt 18919

QY 534 ctgtgtctccagctactctggagagctgaaggtggagagagagagagagagagagagagag 592  
Db 18920 ctgtgtctccagctactctgaagagagagagagagagagagagagagagagagagag 18979

QY 593 agtctacatgacatgattgagctcactgacacagcctgggtgagagagagagagagagag 652  
Db 18980 aggtgtgtgaaactatgtagcgcagctgacacagcctgtgtgga-acaaaaagagagac 19038

QY 653 ctgtctcaaaaaaaataag 707  
Db 19039 ctgtctcagggaaaaaaacaaaaaaacaaataattgtatctgtttgttggaaccttt 19098

QY 708 -----tgaggggag 746  
Db 19099 ttctccacccattctttttcttctgagttcccccaggggtgagagagagagagagagag 19158

QY 747 ggaaggaagaaaaag-----atgaaaaaaagaaaaaaacaga 782  
Db 19159 ttgtctcagagaggtgaggtatctgtctgctcagcacaatacaagtagaaaaatcagtt 19218

QY 783 tgaacag 842  
Db 19219 ttaagcag 842

QY 843 ccaaaaccccaattattgacccaaggttattctttgagagagagagagagagagagagagag 843  
Db 19279 tctttttctttctttctttctttctttctttctttctttctttctttctttctttct 843

QY 903 tgggctgtggcttttagaaagctc-----atctctggcctttctgag 903  
Db 19339 ct 903

Query Match	8.1%	Score	261.6;	DB	4;	Length	18073;
Best Local Similarity	62.1%;	Pred.	No. 6.7e-55;				
Matches	558;	Conservative	0;	Mismatches	309;	Indels	31;
Gaps							8;

QY	365	tgaagaaacaggccggcaccaatgctgtaatgcctgtaataatcccagcaacttgggagcgtga	424
Ddb	1672	TTAATAAAAAAGGTAGGGCGTGGTGCTCATGCTGTAATCCAG--CATTGGGAGGCCGA	1615
QY	425	ggccagaggatcgcttgagtccagatttgagaccagcttgataaacatgcaaaaacc	484
Ddb	1614	GGTGGCGGATCAC--GAGTCAGGATCGACCATCTCGCTAAATGGGGAAAACC	1557
QY	485	tgctctacaaaaaaatacaaaaattagatgggtgtgtggtgcacgttgtgtgccca	544
Ddb	1556	CCTTTTAC-TTCAAAATTCAAAAATACCCGATTGTCGCGGCCCTTTAATCCCA	1498
QY	545	gctacttgggaagctaagttgggagatgccttgagccccagggatccaagtctaca	604
Ddb	1497	GTTTACCGGAGGCTGAGCCGGAGAATTGCATTAACCCCGAGGTGGAGCTTTCAGTGA	1438
QY	605	gccatgatgtgatactgcactccagcctgggttagacagaccaacacctgtccaaaa	664
Ddb	1437	CCCCAAATGCCCCACTGCACCTCGGCCCTGGG--GGACAGATGAGACCCCCTCAAAA	1379
QY	665	aaagaatatgaaa-----gagaaagaagaagaggagagagagagatgagggg	714
Ddb	1378	AAAAAAAAAAAAATACAGTGGTTGTGCAGTTAGGAAAAGGAGAGAGGAGTAGACAGGTGG	1319
QY	715	aggaggagggggggagggaagggaagggaaggaaggaaggaaggaaggaaggaagga	774

INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 246240 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1..246240  
OTHER INFORMATION: /note= "HLA-H.CONTIG"  
US-08-724-394A-20

Query Match 8.0%; Score 257.8; DB 2; Length 246240;  
Best Local Similarity 61.4%; Pred. No. 2.1e-53;  
Matches 545; Conservative 0; Mismatches 312; Indels 30; Gaps 7;  
QY 376 gccgggcaaatggtcctgtaatccacagcacttggagcctgagccacagcagat 435  
Db 172582 GCCGGTGCATGCTCAGCCTGTATCCAGCAGCTTTGGGA-GCTGAGCGGGCAGAC 172524  
QY 436 cgccttgagctccagctggtgagaccagcctggtataacatggcaaacccctgtctctacaa 495  
Db 172523 CAC-----GAGGTCAAGACCACCCCTAGCCGAAATGGTGAACCCCGCTCTATT- 172473  
QY 496 aaaaaatacaaaattagatgggtgtggtgcatgcacctgtgtccagctacttggga 555  
Db 172474 -AAAAATACAAAGTTAGTGGCGGTGGTGCACATCCCTGTAGTCCAGCTACTGGGA 172416  
QY 556 ggctaagtggtgagcctggtgagccagcaggtgcaagtacactgagccatgattgg 615  
Db 172415 GCGTGAGCAGAGAAATGCTTGAACCTGGGAGCGGAGGTGTCAGCGAGCCGAGATTGC 172356  
QY 616 atcactgcactccgctggtgtagacagagcaagccctgtctccaaaaaaagaaatgaa 675  
Db 172355 GCCACTGCACCTCCAGCCTGGTGACAGATGAGACTCCGCTCTAAAAATAAATAATA 172296  
QY 676 agagaagaagaagaagagagagagagagatgagggagagggaggggggaggaag 735  
Db 172295 ATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 172236  
QY 736 gaaggagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaagaaga 795  
Db 172235 CTCCTGCGGTGTTTCAGCGACTTTCAGAGGCCAACAGAGTAAGTTTCCCTTAGCCC 172176  
QY 796 gaaagact----tacgtaaattcctcatcatgtggtgtgcaagtttgaccccccaaaccc 851  
Db 172175 CTACAGGTTCTTATGTTTAAATTTGTTACTCTCATTTAAGACATAAATAAAGTGGCTTCTC 172116  
QY 852 aattattgaccaaggttattcttctgactgagcgaagggtccgctcctctggtgacctg 911  
Db 172115 CATGAGAGATTATTCGATCCATATTTGTTAGATTGGCGGTTTCTCCTTTGATCTC 172056  
QY 912 ggccttagaagatcatctctggtccttctgagatccatcccttctttttattttctt 971  
Db 172055 TACTTCACACTGACCCACATAAAACATC---ACTGCCGTGTTTTTGTGTTGTTGTTG 171999  
QY 972 gacaggaattctctgtcactcaagcctgagtgagtgagtgagtgagtgagtgagtgag 1031  
Db 171998 GAGACGGAGTCTTCTCTGTTGGCCAGCGTGGAGTGAGTGAGTGAGTGAGTGAGTGAGTGAG 171939  
QY 1032 taacctctgctcccggttcaagcattctcctgctcagctccctga-----gat 1083  
Db 171938 CAAGCTCGGCTCCGGATTACGCCATTTCTCTGCTCAGCTCCCTGAGCAGCTGGGAC 171879  
QY 1084 aacagggtccgcacacacatctggtctaatttttttttttttttttttttttttttttt 1143  
Db 171878 TACAGGCCACCCACCAAGCCCGCTAAATTTTGTATTTTGTATAGATACAGGGGTTTCA 171819  
QY 1144 tcatgtggccaggttgggttcgaactcctgacacgtgagtgagtgagtgagtgagtgagtgag 1203  
Db 171818 CTTTGTAAACAGGATGCTTCGATCTCCTGACCT--CGTATGGGCGCCGCTCAGCGCTC 171761

QY 1204 ccaaatgctggattacagcctgagccactgagccactgagccactgagccactgagat 1250  
Db 171760 CCAAGTGTGGATTACAGAGTGAGCCACTGCGCCGCGCCCGTT 171714

## RESULT 7

US-08-724-394A-21/c  
Sequence 21, Application US/08724394A  
Patent No. 5872237  
GENERAL INFORMATION:  
APPLICANT: Feder, John N.  
APPLICANT: Kronmal, Gregory S.  
APPLICANT: Lauer, Peter M.  
APPLICANT: Ruddy, David A.  
APPLICANT: Thomas, Winston  
APPLICANT: Tsuchihashi, Zenta  
APPLICANT: Wolff, Roger K.  
TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el  
TITLE OF INVENTION: Sequences and Antibodies Thereeto  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/724,394A  
FILING DATE: 01-OCT-1996  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitts, Renee A.  
REGISTRATION NUMBER: 35,136  
REFERENCE/DOCKET NUMBER: 017957-000100  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-576-0200  
TELEFAX: 415-576-0300  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 246240 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1..246240  
OTHER INFORMATION: /note= "HLA-H.CONTIG"  
US-08-724-394A-21

Query Match 8.0%; Score 257.8; DB 2; Length 246240;  
Best Local Similarity 61.4%; Pred. No. 2.1e-53;  
Matches 545; Conservative 0; Mismatches 312; Indels 30; Gaps 7;  
QY 376 gccgggcaaatggtcctgtaatccacagcacttggagcctgagccacagcagat 435  
Db 172582 GCCGGTGCATGCTCAGCCTGTATCCAGCAGCTTTGGGA-GCTGAGCGGGCAGAC 172524  
QY 436 cgccttgagctccagcctggtgagaccagcctggtataacatggcaaacccctgtctctacaa 495  
Db 172523 CAC-----GAGGTCAAGACCACCCCTAGCCGAAATGGTGAACCCCGCTCTATT- 172473  
QY 496 aaaaaatacaaaattagatgggtgtggtgagcctggtggtgagcctggtggtgagcctggtgag 555  
Db 172474 -AAAAATACAAAGTTAGTGGCGGTGGTGCACATCCCTGTAGTCCAGCTACTGGGA 172416

QY	556	ggctaagtggtggaggatcgcttgagcccgaggagtcgaagtctacactgagccatgattgg	615
Db	172415	GGCTGAGGCAGGAGAAATGCTTGAACCTCGGAGCGGAGGTTGCAGCGAGCGAGATTGC	172356
QY	616	atcacctgactccagcctgggtagacagagcaagaccctgtctcaaaaaaagaatgaa	675
Db	172355	GCACATGCACTCCAGCGCTGGTGACAGAGTGAGACTCCGCTCAAAATTAATAATAATA	172296
QY	676	agagaagaagaagagagagagagagagatgagggagagagagggggggagggaag	735
Db	172295	ATAATATAATAATAATAATAATAATAATGATGCATTTTATCCTATTATCTTC	172236
QY	736	gaagggaagggaagaaaaaagatgaaaaaagaaaaaacaagatgaaacagagca	795
Db	172235	CTCTGTGCGTGGTTTTTCAGCGACTCTTCAGAGGCCAAAGAGTAAGTTTTCCTTAGGCC	172176
QY	796	gaagagactt----tagtaaatgtctcatcattggtgtcaagttgacccccaaaccc	851
Db	172175	CTACAGGTTCTTATGTTTATTTTGTACTCTCATTTAAGACATAATAATAAGTGGCTTCTC	172116
QY	852	aatttatgaccaggattactttgactgaggaagggggcogctctccctgggccttg	911
Db	172115	CATCAAGATTATTTCTGCATCCATTATTTGGTAAGATTGGCGGTTTTCTCCTTTGATCTC	172056
QY	912	ggotttagaaagctatctctggccttctgagatccatcccttctttttttttttt	971
Db	172055	TACTTTCACATGACCCACATAAAACATC---ACTGCCTGTTTTTTGTTGTTGTTTG	171999
QY	972	gacacggagcttgctctgtcactcaagctggagtgacgtggcatgatctcgaactcagt	1031
Db	171998	GAGACGGAGCTTGTCGTGTGTGCCAGGTGGAGTGCAGTGTGTGTATCTCCGCTCACAG	171939
QY	1032	taacctgtgctcccggtttcaagcgattctctgctcagctcctcga-----gat	1083
Db	171938	CAAGCTCCGGCTCCGGATTACGCCCATTTCTCTGCCTCAGCCTCCTGAGCAGCTGGAC	171879
QY	1084	acagcgccgcgcacacacatctggttaattttttgtatttttagtaaaagactgggttca	1143
Db	171878	TACAGGCACCCACACCAAGCCCGGCTAATTTTGTATTTTATAGTAGATACGGGGTTTCA	171819
QY	1144	tcatgttggccaggttggtttcgaacctctgacctgagctgagctgccacactggcttc	1203
Db	171818	CTTTGTTTAACACAGATGGTCTCGATCTCTGAGCT---CGTGATCGCCCGCCTCAGCCTC	171761
QY	1204	ccaaagtgtgtgattacaggcatgagccactgcgccagctcagat	1250
Db	171760	CCAAAGTGTGGGATTACAGAGTGAAGCCTGCGCCCGGCCGGT	171711

## RESULT 8

US-08-724-394A-22/c  
: Sequence 22, Application US/08724394A

; Patent No. 5872237

GENERAL INFORMATION:

APPLICANT: Feder, John N.

; APPLICANT: Kronmal, Gregory S.

; APPLICANT: Lauer, Peter M.

; APPLICANT: Ruddy, David A.

APPLICANT: Thomas, Winston

; APPLICANT: Tsuchihashi, Zenta

APPLICANT: Wolff, Roger K.

TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el

;	TITLE OF INVENTION:	Sequences and Antibodies Thereof
1.	NUMBER OF SEQUENCES:	21

; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:

CORRESPONDENCE ADDRESS: : ADDRESS: TOWNSEND

ADDRESSEE: TOWNSEND AND TOWNSEND and CREW LLP  
STREET: Two Embarcadero Center, 8th Floor

STREET: TWO EMPEROR CENTER, 600 FROST  
CITY: San Francisco

CITY: SAN FRANCISCO  
STATE: CA

COUNTRY: USA

ZIP: 94111-38

; COMPUTER READABLE

—

```

, MEDIUM TYPE: Floppy disk
, COMPUTER: IBM PC compatible
, OPERATING SYSTEM: PC-DOS/MS-DOS
, SOFTWARE: PatentIn Release #1.0, Version #1.30
, CURRENT APPLICATION DATA:
, APPLICATION NUMBER: US/08/724,394A
, FILING DATE: 01-OCT-1996
, CLASSIFICATION: 536
, ATTORNEY/AGENT INFORMATION:
, NAME: Fitts, Renee A.
, REGISTRATION NUMBER: 35,136
, REFERENCE/DOCKET NUMBER: 017957-000100
, TELECOMMUNICATION INFORMATION:
, TELEPHONE: 415-576-0200
, TELEFAX: 415-576-0300
, INFORMATION FOR SEQ ID NO: 22:
, SEQUENCE CHARACTERISTICS:
, LENGTH: 246240 base pairs
, TYPE: nucleic acid
, STRANDEDNESS: not relevant
, TOPOLOGY: not relevant
, MOLECULE TYPE: cdna
, FEATURE:
, NAME/KEY: misc_feature
, LOCATION: 1..246240
, OTHER INFORMATION: /note= "HLA-H.CONTIG"
US-08-724-394A-22

```

Query Match 8.0%; Score 257.8; DB 2; Length 246240;  
Best Local Similarity 61.4%; Pred. No. 2.1e-53;  
Matches 545; Conservative 0; Mismatches 312; Indels 30;

QY 376 gcgaggcaaatggctaagtgcctgtaatacccagcaacttgggaggtcgagccagaggaat 435

Db 172582 GCGGGGTGCAATGGCTCAGCGCTGTAATCCAGCACTTTGGGA-GCTGAGGCGGCAGAC 172524

QY 436 cgcttgagctccagagattgagaccagctgataacatggcaaacacctgtctctaca 495

Db 172523 CAC-----GAGGTCAAGACCACTTAGCCGAATGGTGAAACCCGCTCTATT- 172473

QY 496 aaaaaatcacaaatttagatgggtggcatgacctgtggtgccagctacttgga 555

Db 172474 -AAAAATACAAAGTTAGCTGGGCGTGGGCATGCCTGTAGTCCAGCTACTGGGA 172416

QY 556 ggctaagtgggaggtatcgcttgagcccgaggagtcagtctacactgagccatgatgg 615

Db 172415 GGCTGAGCGCAGGAATGCTTGAATCTGGAGCGCGAGGTTGCAGCGGACGAGATTGC 172356

QY 616 atcacitgcactccagctgggttagacagagacaagaccctgtctcaaaaaaagaatgaa 675

Db 172355 GCACATGCATCTCAGCGCTGGTGACAGGTGAGACTCCGCTCAAAATATATATAATA 172296

QY 676 agagaagaaaagaagagagagagagagatgaggggagggagggaggggggaggaa 735

Db 172295 ATAAATAATAATAATAATAATAATAATAATGGATCAATTTATCCTATTAACTTC 172236

QY 736 gaaggaggaaggagaaagaaagatgaaaaagaaaaaacaagatgaaacagagga 795

Db 172235 CTCTGTGCGTGGTTTTTCAGCGACTCTTCAGAGCGCCAAAGAGTAAGTTTTCCTTAGCC 172176

QY 796 gaaagactt---tagtgaattgtcatcatgtggtgtcaagtttgaccccaaaccc 851

Db 172175 CTACAGGTTCTTATGTTTAAATTTGTACTCTCATTTTAAGACATAATAAAGTGGCTCTC 172116

QY 852 aattattgaccaaggtattctttgactgaggcaagggtccgctctctctgggacctg 911

Db 172115 CATGAGATTATTTCTGCATCCCAATATTATGGTAAGATTGGCCGTTTCTCCTTGNATCTC 172056

QY 912 ggctttagaagctcatctctggcctttctgagatccatccctttcttttttttt 971

Db 172055 TACTTCACATGCAGCCACATAAACATC---ACTGCCTGTTTTTTGTTGTTGTTGTTG 171999











QY 654 tftctcaaaaaaagaatgaagagaaagaaagaaagagagagagagagagagag 713  
 Db 14536 GTCTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 14477  
 QY 714 gag 773  
 Db 14476 AATTTGAACAGTATTAAGTGAAGTTCCTCTCTCTCTCTCTCTCTCTCT 14417  
 QY 774 aaacaagatga-----acagagagagagagagagagagagagagagag 820  
 Db 14416 CATTCCTAAAGCAGTGTAACTGTGTCTATATATATATATATATATATAT 14357  
 QY 821 catgtgtgtcaagtttgaccccaaaaccccaatttatigacaaagttattct----- 874  
 Db 14356 CATTTTCTGTAGCCCAACAGGTACATGCATATATATATATATATATATAT 14297  
 QY 875 -----ttgactga 882  
 Db 14296 CTGAATCCATGTTTATACATAATTTAGTATCTTTTACACACTCTTCTGTACTGTACTTA 14237  
 QY 883 ggaagaggggtccgctctcctggtggtggtggtggtggtggtggtggtggtggt 942  
 Db 14236 AAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 14177  
 QY 943 agatccatcccttcttcttcttcttcttcttcttcttcttcttcttcttcttct 1002  
 Db 14176 TCATGACTTTTCTCATTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 14117  
 QY 1003 gagtgcagtgcatgctgactcactgtaacctctgctccc-gggttcaagcgtatc 1061  
 Db 14116 GAGTGCAGTGGCGTATCTCCATCACTGTAACTCGCCGCCCCAGAGTTTCAAGCGATTC 14057  
 QY 1062 tctgctcctcagctc-----ctgagataacagagggcgccgcacacacatctggtctaat 1113  
 Db 14056 TCCTGCTCTACGCCCAAGTAGCTGGGATTTACAGGCGCTTGCACACACACCTGGCTAAT 13997  
 QY 1114 ttgtgtatttttagaagctggtttctcatctgttggcaggttgggttgcgaactcct 1173  
 Db 13996 TTTTCCATTTTGTAGAGGAGGTTTCTGCATATTTGCGCAGGCTGCTTGAACCTCT 13937  
 QY 1174 gacctgaggtgagctgccccctggtcctcccaagtgctgggattacagagcatgagcca 1233  
 Db 13936 GACCTCAAGTATCAGCTGCTGCTTGGCTTCCAAAGTGCTAGGATTACAGGCATAAGCCA 13877  
 QY 1234 ctgagccagctcagatccatccctttct 1262  
 Db 13876 CCACACCTGGCCCTCATCTTTATTTT 13848

RESULT 13  
 US-09-210-748A-3  
 ; Sequence 3, Application US/09210748A  
 ; Patent No. 6335156  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hermeking, Heiko  
 ; APPLICANT: Vogelstein, Bert  
 ; APPLICANT: Kinzler, Kenneth  
 ; TITLE OF INVENTION: 14-3-3 SIGMA ARREST THE CELL CYCLE  
 ; FILE REFERENCE: 1107.77810  
 ; CURRENT APPLICATION NUMBER: US/09/210,748A  
 ; CURRENT FILING DATE: 1998-12-15  
 ; PRIOR APPLICATION NUMBER: 60/069,416  
 ; PRIOR FILING DATE: 1997-12-18  
 ; NUMBER OF SEQ ID NOS: 18  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 3  
 ; LENGTH: 7680  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-210-748A-3

Query Match

7.8%; Score 251.4; DB 4; Length 7680;

Best Local Similarity 61.0%; Pred. No. 1.5e-52;  
 Matches 575; Conservative 0; Mismatches 286; Indels 82; Gaps 7;  
 QY 371 aacagcgggcacaaatgctgaatgctgaatccacacacttttgagaggtgagccag 430  
 Db 5422 atctgcccggcatggtgcttatgctataatttcacactttggaaggtcgagcagg 5481  
 QY 431 aggatcgtctgagctccagagttgagaccagcctggataacatgcaaaacccctgtctc 490  
 Db 5482 tggatcacttgagctcagctgtcgagaccagcctggcaacatggtgaacacacatctc 5541  
 QY 491 tac-----aaaaaatacaaaatagatggtgtggtggtggtggtggtggtggtggt 542  
 Db 5542 tacttaaaaaaaatacaaaatagctggtgtggtggtggtggtggtggtggtggtggt 5601  
 QY 543 cagctacttgagagtaaggtggtgagatcgcttgagccagggaggtcaagctcacact 602  
 Db 5602 cagctactggtgaggtggtgaggaatcacttgaaacagggaggtggtggtggtggt 5661  
 QY 603 gagccatgattgatacactgcactccagcctgggtgagacagagcagacccgtgtctcaaa 662  
 Db 5662 gagctgagattgctcactacactccagcctgggtgacaagagcgaactcgtctcaaa 5721  
 QY 663 aaaaagaaatgaaagagaaagaaagagagagagagagagagagagagagagagag 722  
 Db 5722 aaaaagaaataataataatagttg-----ccat 5754  
 QY 723 gggggggag 782  
 Db 5755 ccattctactgtcttccattaaactcgtgtaactcacaagtcoccatcttatagttac 5814  
 QY 783 tgaacacagagcagaaagactttagtaaatgtctcatcatctggtgtgcaagtttgacc 842  
 Db 5815 aggaactggtggttcacagaggttaataactacttggtgcca-----agg 5854  
 QY 843 ccaaaacccaatttattgaccaaggttattctttgactgaggaaggggtgcgtctctcc 902  
 Db 5855 ccacaaacagctataagaattacatttagcagctgattcccaagataactagctctattc 5914  
 QY 903 tgggctggtggtttagaagctcatctctggtcttctgagatccatccctttctttt 962  
 Db 5915 tgt-----atctcatagacaaacacacacacacacacacacacacacacacac 5958  
 QY 963 attttcttgacacgagctctgctctgctcactcaggtggtgagtcagtgagtcacatc 1022  
 Db 5959 ttttgtttgagacgagctctgctctgctcactccagcgtgagtcagtcagtcacatc 6018  
 QY 1023 gactcactgtaacctctgctcccggttcaagcgtattctctgctcagctc----- 1076  
 Db 6019 ggtcactgcaacgtccgctcccggttcaagcgtattctctgctcagctcagctcccgagt 6078  
 QY 1077 --ctgagataacagcgcgcgcacacacacacacacacacacacacacacacacac 1133  
 Db 6079 agctgggactacagcatgtgccacacacacacacacacacacacacacacacacac 6138  
 QY 1134 ctgggtttcatcatgttggccaggttgggtttcgaactcctgacactgaggtgagctgccca 1193  
 Db 6139 cagggtttctggtttagccagaatggctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 6196  
 QY 1194 ccttgccctcccaagtgctggattacagcgtgagcactgcgcacagctcagatccca 1253  
 Db 6197 cctcagcctcccaagtgctgagatgacagcgtgagcagcgtcgtcgtcgtcgtcgtcgtcgt 6256  
 QY 1254 tccctttctaagggcaaacagttccatggtgcaaaaggggcatg 1296  
 Db 6257 ctattataaattggagagaataagaaaaatacaaaaggggccagg 6299

RESULT 14  
 US-08-257-963B-10  
 ; Sequence 10, Application US/08257963B  
 ; Patent No. 5840686  
 ; GENERAL INFORMATION:







GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 19, 2002, 23:20:20 ; Search time 3895 Seconds  
(without alignments)  
11196.062 Million cell updates/sec

Title: US-09-846-456-1  
Perfect score: 3231

Sequence:

1 acagggcatggfaggcagggtg.....gccacacatccccaccactt 3231

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	336	10.4	763	9	AU121731
C	237.8	7.4	289	9	AW748338
3	221.8	6.9	736	9	AU135588
C	221.2	6.8	457	9	AW816516
C	219.4	6.8	577	12	AQ265389
6	218.6	6.8	544	12	AQ418551
C	215.2	6.7	719	12	AQ475181
C	214.2	6.6	581	12	AQ347610
C	213.4	6.6	338	9	AW023111
10	213.4	6.6	666	12	AG075934
11	213.2	6.6	493	12	B50449
C	213.2	6.6	622	12	AQ537948
13	213	6.6	514	9	AI754653
C	213	6.6	739	12	AQ200209
C	213	6.6	827	9	AV755654
C	212.8	6.6	484	12	AQ314669
C	212.8	6.6	487	12	AQ414845

18	212.4	6.6	501	12	AQ040260
19	212.2	6.6	416	9	AW341978
C	212.2	6.6	448	12	AQ634562
C	212.2	6.6	521	9	AW970571
C	212.2	6.6	521	9	AW970571
C	212.2	6.6	575	12	AQ587429
C	211.4	6.5	577	9	AW973181
C	211.4	6.5	577	9	AW973181
C	211.4	6.5	678	12	AQ387027
C	211.2	6.5	864	10	BG682030
C	211.2	6.5	1005	9	AL524675
C	210.8	6.5	458	9	AI733856
27	210.6	6.5	388	9	AW069227
28	210.6	6.5	437	9	AA644090
C	210.6	6.5	605	12	AQ412598
C	210.4	6.5	584	12	AG102895
31	210.4	6.5	924	10	BG697612
32	210.2	6.5	477	12	AQ221138
C	210	6.5	577	9	AW19392
C	209.8	6.5	541	9	AW973187
C	209.8	6.5	671	12	AG050562
C	209.6	6.5	657	12	AG036967
37	209.4	6.5	461	12	B65075
C	209.4	6.5	642	12	B59854
C	209.4	6.5	832	10	BG697217
C	209.4	6.5	483	9	AI275982
41	209	6.5	490	12	AQ426532
C	209	6.5	616	12	AQ554309
C	209	6.5	1750	11	AF119908
C	208.8	6.5	624	10	BF854308

## ALIGNMENTS

RESULT 1

AU121731

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AU121731 763 bp mRNA linear EST 19-OCT-2000  
AU121731 MAMMA1 Homo sapiens cDNA clone MAMMA1000851 5', mRNA  
sequence.

AU121731 GI:10936966  
EST.  
human.  
HRI human cDNA project  
Unpublished (2000)  
Contact: fakao isogai  
Genomics Laboratory  
Helix Research Institute  
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
Tel: 81-438-52-3951  
Fax: 81-438-52-3952  
Email: genomics@hri.co.jp

HRI human cDNA project; 5' - & 3'-end one pass sequencing: Helix  
Research Institute; cDNA library construction: Department of  
Virology, Institute of Medical Science, University of Tokyo, and  
Helix Research Institute.  
Location/Qualifiers  
1. 763  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="MAMMA1000851"  
/clone\_lib="MAMMA1"  
/tissue\_type="mammary gland"  
/note="Vector: pME18SFL3"

137 a 205 c 260 g 158 t 3 others  
BASE COUNT  
ORIGIN

137 a 205 c 260 g 158 t 3 others  
BASE COUNT  
ORIGIN

137 a 205 c 260 g 158 t 3 others  
BASE COUNT  
ORIGIN



BASE COUNT 163 a 199 c 199 g 170 t 5 others  
ORIGIN

Query Match 6.9%; Score 221.8; DB 9; Length 736;  
Best Local Similarity 99.1%; Pred. No. 6.8e-20;  
Matches 223; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 2894 gtaattgagcgagagtgagtgaggccgagaccgagaccgagaccgaccttctctc 2953  
Db 1 GTAGTGGGAGCGAGAGTGTGGGCGCGGACCGAGACCGAGCGGCTCTCTC 60  
QY 2954 ccgggctcggcgagcgagggcgaggctccgcgcacacacagagccggttctcagggc 3013  
Db 61 CCGGGCTCGGCGAGCGAGCGGGGAGCTCCGCCACCAACAGAGCGGTTCTCAGGGC 120  
QY 3014 gcttgctcctgtttttcccggttctgttttctccttctcggaaagcttgtaa 3073  
Db 121 GCTTTCCTCTGTTTTCCTCCCGGTTCTGTTTCTCCCTCTCCCGAAGGCTGTCAA 180  
QY 3074 ggggtagagaaagagacgcaacacaaaagtggaaaacaggttaa 3118  
Db 181 GGGGTAGGAGAAAGAGACGCAACACAAAGTGGAAACAGTTAA 225

RESULT 4  
AW816516/c  
LOCUS QV0-ST0236-171299-075-b12 ST0236 Homo sapiens cDNA, mRNA sequence.  
DEFINITION  
ACCESSION AW816516  
VERSION AW816516.1 GI:7909510  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 457)  
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.K., Reis,L.F., de Souza,S.J. and Simpson,A.J.  
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
MEDLINE 20202663  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2-QV0-ST0236-171299-075-b12&t3=1999-12-17&t4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 457.

FEATURES  
source  
1. 457  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="ST0236"  
/dev\_stage="Adult"  
/note="Organ: stomach; Vector: puc18; Site:1: SmaI;  
Site:2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under

BASE COUNT 157 a 95 c 115 g 90 t  
ORIGIN

Query Match 6.8%; Score 221.2; DB 9; Length 457;  
Best Local Similarity 85.3%; Pred. No. 1e-19;  
Matches 261; Conservative 0; Mismatches 38; Indels 7; Gaps 1;  
QY 955 ttcttttattttcttgacacgaggtctgtctgtcactcaggtggagtgcaagtgc 1014  
Db 457 TTTTCTTTTCTTTTGTGAGACAGAGTCTTGTCTGCTCCAGGCTGGAGTACAGTGC 398  
QY 1015 atgctctcgactcaactgaacctctgctcccggttcaagcattctctcgtcctcagcc 1074  
Db 397 GTGATCTCAACTCACTCAACCTCTGCTCCGCGTTCAAGCGATTCTCTGCTCAGCC 338  
QY 1075 tcttga-----gataacagggcgccgcacacacatctggtaattttgtatttttag 1127  
Db 337 TCTTGAGTATCTGGACTACATGCACACACACACACCTGGCTAATTTTGTATTTTAG 278  
QY 1128 taaagactgggtttcatcatgttgccaggttggtttcgaactcctgacctgaggtgagc 1187  
Db 277 TAGACACGGGGTTTCAACACGTTTGGCCACGCTGCTCTCGAACTCCTGACCTCAGGTGATC 218  
QY 1188 tgccacacttggtctcccaagtgctggattacagggcatgagccactgcgccagctca 1247  
Db 217 TGCCACCTCAGCCTCCAAAAGTGTGGGATTACAGCGGTGAGCCACTGTGCCCGGCTCT 158  
QY 1248 gatcca 1253  
Db 157 TCTTCA 152

RESULT 5  
AQ265389/c  
LOCUS CITBI-EI-2509010.TF CITBI-EI Homo sapiens genomic clone 2509010,  
DEFINITION DNA sequence.  
ACCESSION AQ265389  
VERSION AQ265389.1 GI:3793589  
KEYWORDS GSS.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 577)  
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and Venter,J.C.  
TITLE Use of a random human BAC End Sequence Database for Sequence-Ready Map Building  
JOURNAL Unpublished (1998)  
COMMENT Other\_GSSs: CITBI-EI-2509010.TR  
Contact: Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: mdadams@tigr.org  
Clones are available from Research Genetics (info@resgen.com). BAC end search page:  
http://www.tigr.org/tldb/humgen/bac\_end\_search/bac\_end\_search.html.  
Seq primer: M13-21  
Class: BAC ends.

FEATURES  
source  
1. 577  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="2509010"  
/clone\_lib="CITBI-EI"  
/sex="male"





Fax: 301 838 0208  
 Email: hbe@tigr.org  
 Clones are available from Research Genetics (info@resgen.com). BAC  
 end search page:  
 http://www.tigr.org/tdb/humgen/bac\_end\_search/bac\_end\_search.html.  
 Seq primer: M13-21  
 Class: BAC ends.

# FEATURES

source  
 1. 719  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="2589B9"  
 /clone\_lib="CITBI-El"  
 /sex="male"  
 /cell\_type="sperm"  
 /note="Vector: pBelobAC11; Site\_1: EcoRI; Site\_2: EcoRI;  
 Caltech Human BAC Library D"  
 BASE COUNT 228 a 162 c 180 g 149 t  
 ORIGIN

Query Match 6.7%; Score 215.2; DB 12; Length 719;  
 Best Local Similarity 75.7%; Pred. No. 5e-19;  
 Matches 283; Conservative 0; Mismatches 83; Indels 8; Gaps 1;  
 QY 893 tcgcgtctctggcctggcctttagaaagctcatctctggccttcttgagatccatcc 952  
 Db 431 TTCCATCTCTAGAGCCCTGCGGCAACTGGTTTATCTCTGATAAATGGATCCATCTTT 372  
 QY 953 cttcttt 1012  
 Db 371 TTTTGTCTTTTGTATTTTGTAGACAGATCTTGTCTGTCACTCAAGCTGGAGTGCAGTG 312  
 QY 1013 gcatgatctcagactgaactgttaacctgtctcccccgggttcaagcgattctctgcctcag 1072  
 Db 311 GCATGCTCTCGCTCACTGCAACCTCCGCCCTCCGGGTTCAAGCGATTCTCCTGCCCTCAG 252  
 QY 1073 cctc-----ctgagataacagcgcccccaccacatctgctgaatttttttttttt 1124  
 Db 251 CTCCTCCGGAACAGCTGGGATTACAGGCACCTGCCACCATGCTGCTGATTTTGTATTTT 192  
 QY 1125 tagtaaacactgggttttctcatatgttgccaggtgtgttttcgaactcctgcactgaggtg 1184  
 Db 191 TAGTAGAGACGGGGTTTGGCCATGTTGCCAGCTGCTCGAACTCTGACCTCAAGTG 132  
 QY 1185 agctgccacacttgctcccaagtgctggattacaggcatgagcaactgcgccacg 1244  
 Db 131 AFCCCTCTCGCTCGCCCTCCCAAGTGTGGGATTGTCAGGTGTGAGCCACCGCGCGCAGT 72  
 QY 1245 teagatccatccct 1258  
 Db 71 CTGGATCCATCTT 58

RESULT 8  
 AQ347610/c  
 LOCUS  
 DEFINITION  
 DNA sequence.  
 ACCESSION  
 AQ347610  
 VERSION  
 AQ347610.1 GI:4172506  
 KEYWORDS  
 GSS.  
 SOURCE  
 human.  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 1 (bases 1 to 581)  
 AUTHORS  
 Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter  
 J.C.  
 TITLE  
 Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready  
 Map Building  
 JOURNAL  
 Unpublished (1997)  
 COMMENT  
 Other\_GSSs: RPCI11-126022.TV

Contact: Shaying Zhao, William Nierman, Mark Adams  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: hbe@tigr.org  
 Clones are derived from the human BAC library RPCI-11. For BAC  
 library availability, please contact pieter de Jong  
 (pieter@dejong.med.buffalo.edu). Clones may be purchased from  
 BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from  
 Research Genetics (info@resgen.com). BAC end search page:  
 http://www.tigr.org/tdb/humgen/bac\_end\_search/bac\_end\_search.html  
 Seq primer: SP6  
 Class: BAC ends.

# FEATURES

source  
 1. 581  
 /organism="Homo sapiens"  
 /db\_xref="GDB:7548357"  
 /db\_xref="taxon:9606"  
 /clone="RPCI-11-126022"  
 /clone\_lib="RPCI-11"  
 /sex="Male"  
 /cell\_type="Lymphocytes"  
 /note="Vector: pBACE3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
 RPCI11 Human Male BAC Library"  
 BASE COUNT 180 a 138 c 130 g 133 t  
 ORIGIN

Query Match 6.6%; Score 214.2; DB 12; Length 581;  
 Best Local Similarity 81.2%; Pred. No. 7.4e-19;  
 Matches 264; Conservative 0; Mismatches 53; Indels 8; Gaps 1;  
 QY 954 ttcttt 1013  
 Db 454 TTTCTTCTATTTATTTTGTAGACAGATCTTGTCTGTACCCAGGTGAGTGCAGTG 395  
 QY 1014 catgatctgaactgaacctctgcctcccggttcaagcgattctctgcctcagc 1073  
 Db 394 CTGGATCTCGGCTCACTGCAAACTCTGCCCTCTGGTTCAAGCGATTCTCCTGCCCTCAGC 335  
 QY 1074 ctc-----ctgagataacagcgcccccaccacatctgctgaatttttttttttt 1125  
 Db 334 CTCGAGTAGTCTGGGATTACAGCGCGCTGCCATCATGCTGCTGCTATTTTGTATTTT 275  
 QY 1126 agtaaacactgggtttctcatatgttgccaggtgtgttttcgaactcctgcactgaggtga 1185  
 Db 274 AGTAGAGATGGGGTTTTCACCATCTGTCAGAGATTGGTCTCGAACTCTGACCTCAGGTGA 215  
 QY 1186 gctgccacacttgccctcccaagtgctgggattacaggcatgagcaactgcgccacgt 1245  
 Db 214 TCTGCCACCTTGGCCCTCCCAAGTGTGGGATTACAGCGGTGAGCCACCGCCAGCC 155  
 QY 1246 cagatccatccctttctaaagggcaa 1270  
 Db 154 TTTATTTCTTTTCTTAGCTGAAA 130

RESULT 9  
 AW023111/c  
 LOCUS  
 DEFINITION  
 IMAGE:2486822 5', mRNA sequence.  
 ACCESSION  
 AW023111  
 VERSION  
 AW023111.1 GI:5876641  
 KEYWORDS  
 EST.  
 SOURCE  
 human.  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 1 (bases 1 to 338)  
 AUTHORS  
 Robertson,N.G., Khetarpal,U., Gutierrez-Espeleta,G.A., Bieber,F.R.

TITLE Isolation of novel and known genes from a human fetal cochlear cDNA library using subtractive hybridization and differential screening

JOURNAL Genomics 23, 42-50 (1994)

MEDLINE 95130111

COMMENT Contact: Morton, C. C.  
Departments of Pathology and Obstetrics, Gynecology and Reproductive Biology  
Brigham and Women's Hospital  
75 Francis Street, Harvard Medical School, Boston, MA 02115, USA  
Tel: 617 732 7980  
Fax: 617 738 6996  
Email: cmorton@rics.bwh.harvard.edu  
DNA sequencing and analyses were performed by National Institutes of Health Intramural Sequencing Center (NISC; see <http://www.nisc.nih.gov>).  
This clone is available royalty-free through LNHL; contact the IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
Plate: L246M188 row: G column: 15  
Seq primer: M13RP1 reverse primer (ABI).

FEATURES	SOURCE
1. <i>General</i>	
2. <i>Specific</i>	
3. <i>Other</i>	
4. <i>Other</i>	
5. <i>Other</i>	
6. <i>Other</i>	
7. <i>Other</i>	
8. <i>Other</i>	
9. <i>Other</i>	
10. <i>Other</i>	
11. <i>Other</i>	
12. <i>Other</i>	
13. <i>Other</i>	
14. <i>Other</i>	
15. <i>Other</i>	
16. <i>Other</i>	
17. <i>Other</i>	
18. <i>Other</i>	
19. <i>Other</i>	
20. <i>Other</i>	
21. <i>Other</i>	
22. <i>Other</i>	
23. <i>Other</i>	
24. <i>Other</i>	
25. <i>Other</i>	
26. <i>Other</i>	
27. <i>Other</i>	
28. <i>Other</i>	
29. <i>Other</i>	
30. <i>Other</i>	
31. <i>Other</i>	
32. <i>Other</i>	
33. <i>Other</i>	
34. <i>Other</i>	
35. <i>Other</i>	
36. <i>Other</i>	
37. <i>Other</i>	
38. <i>Other</i>	
39. <i>Other</i>	
40. <i>Other</i>	
41. <i>Other</i>	
42. <i>Other</i>	
43. <i>Other</i>	
44. <i>Other</i>	
45. <i>Other</i>	
46. <i>Other</i>	
47. <i>Other</i>	
48. <i>Other</i>	
49. <i>Other</i>	
50. <i>Other</i>	
51. <i>Other</i>	
52. <i>Other</i>	
53. <i>Other</i>	
54. <i>Other</i>	
55. <i>Other</i>	
56. <i>Other</i>	
57. <i>Other</i>	
58. <i>Other</i>	
59. <i>Other</i>	
60. <i>Other</i>	
61. <i>Other</i>	
62. <i>Other</i>	
63. <i>Other</i>	
64. <i>Other</i>	
65. <i>Other</i>	
66. <i>Other</i>	
67. <i>Other</i>	
68. <i>Other</i>	
69. <i>Other</i>	
70. <i>Other</i>	
71. <i>Other</i>	
72. <i>Other</i>	
73. <i>Other</i>	
74. <i>Other</i>	
75. <i>Other</i>	
76. <i>Other</i>	
77. <i>Other</i>	
78. <i>Other</i>	
79. <i>Other</i>	
80. <i>Other</i>	
81. <i>Other</i>	
82. <i>Other</i>	
83. <i>Other</i>	
84. <i>Other</i>	
85. <i>Other</i>	
86. <i>Other</i>	
87. <i>Other</i>	
88. <i>Other</i>	
89. <i>Other</i>	
90. <i>Other</i>	
91. <i>Other</i>	
92. <i>Other</i>	
93. <i>Other</i>	
94. <i>Other</i>	
95. <i>Other</i>	
96. <i>Other</i>	
97. <i>Other</i>	
98. <i>Other</i>	
99. <i>Other</i>	
100. <i>Other</i>	

```

1. 338
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2486822"
/tissue_lib="Morton Petal Cochlea"
/tissue_type="cochlea"
/dev_stage="16-22 week fetus"
/lab_host="504R cells (kanamycin
/note="Organ: ear; vector: pBlues
Site: 2; XhoI; Reference: Genomics
unidirectionally. Primer: Oligo d
37% of inserts <0.5 kb, 56% 0.5-1
XR vector. Library constructed by
-5' adaptor sequence: 5' GAATTCGG
sequence: 5' CTCGAGTTTTTTTTTTT
80 c 99 g 59 t

```

Query Match 6.6%; Score 213.4; DB 9; Length 338;  
Best Local Similarity 83.9%; Pred. No. 1.2e-18;  
Matches 256; Conservative 0; Mismatches 41; Indels 8

QY	954	ttctcttttattttctcttgacacggagctctgtctgtcactcagcgctgagtgcaagtgg	1013
Db	317		
QY	1014	catgatctcgactcaactgaacctctgccctcccggttcaagcgattctcgtgcctcagc	1073
Db	257	CGTGATCTGGCTCACTGCACACTTGCCTCCCTGGGTTCAAGTGATCTCCCTGCCTCAGC	198
QY	1074	ctc-----ctgagataacaggcgcccgccaccacatctcgtgctaattttgtattttt	1125
Db	197	CTCCCGAGTAGCTGAGATTACAGGTGCCACACCATGCCGCTAAATTTTGTATTTT	138
QY	1126	agtaaaagactgggtttcatcatgttggtccaggttggtctgcgaactcctcgaactgaggtga	1185
Db	137	AGTAAAGACGGGGTTTCAACCATGTGGCCAGCGTGTCTGTGAACCTCCTGACCTCAAGTGA	78
QY	1186	gctgcgcaacctggcctcccaagctcgtaggattacaggcattagcccaactgogcccaagt	1245
Db	77	TCCGCCCGCCTCGGCCCTCCCAAGTGGTGGATTACAGGCATGAACCAACACGCTTAGCC	18
QY	1246	cgagt	1250
Db	17	CAGAT	13

RESULT 10  
AG075934

LOCUS	AG075934	666 bp	DNA	linear	GSS 03-NOV-2001
DEFINITION	Pan troglodytes	DNA, clone: PTB-068P19.R,	genomic	survey sequence.	

ACCESSION VERSION KEYWORDS SOURCE	ORGANISM	REFERENCE AUTHORS	TITLE JOURNAL REFERENCE AUTHORS	TITLE JOURNAL
--	----------	----------------------	--	------------------

**COMMENT**

Sequencing: M13Rev

Vector : pKS145  
R.Site 1 : SacI  
R.Site 2 : SacI.

FEATURES	SOURCE
1. <i>Age</i>	1. <i>Age</i>
2. <i>Gender</i>	2. <i>Gender</i>
3. <i>Marital Status</i>	3. <i>Marital Status</i>
4. <i>Education</i>	4. <i>Education</i>
5. <i>Income</i>	5. <i>Income</i>
6. <i>Occupation</i>	6. <i>Occupation</i>
7. <i>Religion</i>	7. <i>Religion</i>
8. <i>Political Affiliation</i>	8. <i>Political Affiliation</i>
9. <i>Health Status</i>	9. <i>Health Status</i>
10. <i>Travel History</i>	10. <i>Travel History</i>
11. <i>Employment Status</i>	11. <i>Employment Status</i>
12. <i>Family Size</i>	12. <i>Family Size</i>
13. <i>Home Ownership</i>	13. <i>Home Ownership</i>
14. <i>Vehicle Ownership</i>	14. <i>Vehicle Ownership</i>
15. <i>Insurance Status</i>	15. <i>Insurance Status</i>
16. <i>Charitable Contributions</i>	16. <i>Charitable Contributions</i>
17. <i>Volunteer Work</i>	17. <i>Volunteer Work</i>
18. <i>Substance Use</i>	18. <i>Substance Use</i>
19. <i>Mental Health</i>	19. <i>Mental Health</i>
20. <i>Life Satisfaction</i>	20. <i>Life Satisfaction</i>

```
1. .666
/db_xref="Pan troglodytes"
/ob_xref="taxon:9598"
/clone="PTB-068P19_R"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee
```

BASE COUNT	240 a	120 c	190 g	116 t
ORIGIN				

Query Match	6.6%;	Score 213.4;	DB 12;	Length 666;
Best Local Similarity	69.7%;	Pred. No. 8.8e-19;		
Matches 317; Conservative	0;	Mismatches 136;	Indels 2;	

QY	347	ttctctccattttaagaatgaagaacacagccggccacaatggctaatgctgtaatccc	406
Db	81	TGCTTTCTTTTAAAGAACCAAGTGGAGGTGAGCTGAGCACAGTGGCTTTACACTTGTAAATCCC	140
QY	407	agcaccttgggagcctgagccagagatcgccttgagctccagagtttgaagaccacctg	466
Db	141	AGCACTTTGGGAGCCGAGGCATGAGATGCGTTTGAGCCCGGAGTTCGAGACCAACCTTG	200
QY	467	gatacatggcaaacacctgtctctcacaaaaataacaaaaattagatgggtgtggtgg	526
Db	201	GGCAACATATGAAACTGTCTCGGCAAAAAATAAAAAAATTAAGCTGAGCCTGGTAG	260
QY	527	catgcacctgtgtcccaactacttggaggcctaagttggaggatcgcctgaagccagg	586
Db	261	CACACACCTGTAGTCCACGCTACTCAGGAGGCTGAGGTTGGAGGGTCACTTGAGCCCGAG	320
QY	587	gagtcacgtctacactgagccatgattggtacactgcactccagcctgggtagacagac	646
Db	321	AGTTTAAAGCTGTAGTGAGCTGTGATCATACCATTGCACCTCCAGCCTGAGCA-ACAGAGC	379
QY	647	aagaccctgtctcaaaaaaataatgaagagagaagaagaagaggagagagaggaag	706
Db	380	AAAGCTTGTCAGAAAGAGAAAGGGGAGAGA-AGAAAGGGAGAGAGAAAGTTAAAGAGAGA	438
QY	707	atgaggggagggagggggggaaggaaggaaggaaggaaggaaggaaggaagatgaaa	766





	Query Match	6.6%	Score 213;	DB 9;	Length 827;
	Best Local Similarity	84.2%;	Pred. No. 9e-19;		
	Matches 255;	Conservative	40;	Mismatches	Indels 8; Gaps 1;
Qy	954	t t c t t t t a t t t c t g a c a c g a g t c t g c t c t c a c t c a g c t g g a g t g c a t g g	1013		
Db	369	t r i t	310		
Qy	1014	c a t g a t c t c g a c t c a c t g a a c c t g t c c t c o c g g t t c a a g c g a t t c t c t g c t c a g c	1073		



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 03:11:50 ; Search time 5225.75 Seconds  
(without alignments)  
1429.608 Million cell updates/sec

Title: US-09-846-456-2

Perfect score: 357

Sequence: 1 tggaggctcagctgagagg.....gagggaaggaagctgtgttg 357

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

GenEmbl: *
1: gb_ba: *
2: gb_htg: *
3: gb_in: *
4: gb_ov: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_on: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vi: *
30: em_htg_hum: *
31: em_htg_inv: *
32: em_htg_other: *
33: em_hgo_inv: *

```

# ALIGNMENTS

RESULT	1	AX351030	357 bp	DNA	linear	PAT 06-FEB-2002
AX351030	AX351030	Sequence 2 from Patent WO0183746.				
LOCUS	AX351030	Sequence 2 from Patent WO0183746.				
DEFINITION	AX351030	Sequence 2 from Patent WO0183746.				
ACCESSION	AX351030	Sequence 2 from Patent WO0183746.				
VERSION	AX351030.1	GI:18616386				
KEYWORDS		human.				
SOURCE		human.				
ORGANISM		Homo sapiens				
REFERENCE		Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
TITLE		1 (sites)				
JOURNAL		Rosier-Montus, M.F., Prades, C., Lemoine, C., Naudin, L., Deneffe, P.,				
FEATURES		Brewer, B., Duverger, N., Remaley, A. and Santamarina-Fojo, S.				
source		Regulatory nucleic acid sequences of the abcl gene				
source		Patent: WO 0183746-A 2 08-NOV-2001;				
source		Aventis Pharma S.A. (FR)				
source		Location/Qualifiers				
source		1. .357				
BASE COUNT		64 a 78 c 116 g 99 t				
ORIGIN						

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB, ID	Description
------------	-------	-------	--------	--------	-------------







QY 1 tggaggtctcagctgagagggctgattagcagctcctcattgtgtatggtttgcagca 60  
 Db 53130 TGGAGGTCTCAGCTGAGAGGGCTGGATTAGCAGCTCCTCATTTGGTGTATGGTTTGCAGCA 53189  
 QY 61 ataactgatggctgtttccctcctcctgctttatcttcttaagtttaataaccagcagggcgct 120  
 Db 53190 ATAACGTATGGCTGTTCCTCCCTCCCTGCTTTATCTTCAAGTTAATGACACGACACNGGCGT 53249  
 QY 121 ccctgctgcagctcctgctgctcctccagggctcccgagccacacgctgggctgct 180  
 Db 53250 CCTGCTGTGAGCTCTGGCGCTGCTTCCAGGGCTCCGAGGCACACGCTGGGGTGTCT 53309  
 QY 181 gctcagggaaacatggcattgttggcctcagctgaggttgcctgctgctggaagaacactact 240  
 Db 53310 GGCTGAGGAACATGGCTTGTGGCTCAGCTGAGGTTGCTGCTGTGGAAGAACCTCACT 53369  
 QY 241 ttcaagaagaacacagtaagcttggtttttcagcagcgggggttctctcattttt 300  
 Db 53370 TTCAGNAGNAGAACAGTAAGCTTGGGTTTTTTCAGCAGCGGGGGGTCTCTCATTTTTT 53429  
 QY 301 tcttgggttttggattgggagggagggagggagggaggaagagctgtgttg 357  
 Db 53430 TCTTTGTGTTTTGATTGGGGATTGGAGGAGGAGGAGGAGGAAGAGCTGTGTG 53486  
 RESULT 4  
 HSA252277 HSA252277 480 bp DNA linear PRI 08-JAN-2001  
 LOCUS Homo sapiens partial ABC-1 gene for ATP-binding cassette transporter-1, exon 2.  
 DEFINITION  
 ACCESSION AJ252277  
 VERSION ABC-1 gene; ATP-binding cassette transporter-1.  
 KEYWORDS human.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 480)  
 AUTHORS Porsch-Oezcuuermez,M., Langmann,T. and Schmitz,G.  
 TITLE Cloning and Characterization of the human ATP-binding Cassette Transporter-1 (ABC-1) Promoter  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 480)  
 AUTHORS Porsch-Oezcuuermez,M.K.  
 TITLE Direct Submission  
 JOURNAL Submitted (07-JAN-2000) Porsch-Oezcuuermez M.K., Institute for Clinical Chemistry, University of Regensburg, Franz-Josef-Strauss-Allee 11, 93042 Regensburg, GERMANY  
 FEATURES  
 source  
 1..480  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="9"  
 /map="9q31"  
 /cell\_type="leukocyte"  
 /number=189..346  
 /gene="ABC-1"  
 /number=189..280  
 /gene="ABC-1"  
 /function="cholesterol efflux regulatory protein"  
 /number=189..346  
 /gene="ABC-1"  
 /number=281..346  
 /gene="ABC-1"  
 /function="cholesterol efflux regulatory protein"  
 /codon\_start=1  
 /product="ATP-binding cassette transporter-1"  
 /protein\_id="CAC21428.1"  
 /db\_xref="GI:12140345"  
 /translation="MACWPQLRLLLKLNLFRRRQT"  
 89 a 102 c 155 g 134 t

Query Match 95.3%; Score 340.2; DB 9; Length 480;  
 Best Local Similarity 98.9%; Pred. No. 2.9e-93;  
 Matches 353; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
 QY 1 tggaggtctcagctgagagggctgattagcagctcctcattgtgtatggtttgcagca 60  
 Db 90 TGGAGGTCTCAGCTGAGAGGGCTGGATTAGCAGCTCCTCATTTGGTGTATGGTTTGCAGCA 149  
 QY 61 ataactgatggctgtttccctcctcctgctttatcttcttaagtttaataaccagcagggcgct 120  
 Db 150 ATAACGTATGGCTGTTCCTCCCTCCCTGCTTTATCTTCAAGTTAATGACACGACCTCACT 208  
 QY 121 ccctgctgcagctcctgctgctcctccagggctcccgagccacacgctgggctgct 180  
 Db 209 CCTGCTGTGAGCTCTGGCGCTGCTTCCAGGGCTCCGAGGCACACGCTGGGGTGTCT 268  
 QY 181 ggtcagggaaacatggcattgttggcctcagctgaggttgcctgctgctggaagaacactact 240  
 Db 269 GGCTGAGGGAAACATGGCTTGTGGCTCAGCTGAGGTTGCTGTGTGGAAGAACCTCACT 328  
 QY 241 ttcaagaagaacacagtaagcttggtttttcagcagcgggggttctctcattttt 300  
 Db 329 TTCAGNAGNAGAACAGTAAGCTTGGGTTTTTTCAGCAGCGGGGGGTCTCTCATTTTTT 388  
 QY 301 tcttgggttttggattgggagggagggagggagggaggaagagctgtgttg 357  
 Db 389 TCFTTGTGTTTTGATTGGGGATTGGAGGAGGAGGAGGAGGAAGAGCTGTGTG 445  
 RESULT 5  
 AL353685/c AL353685 129608 bp DNA linear PRI 01-JUN-2001  
 LOCUS Human DNA sequence from clone RP11-31J20 on chromosome 9, complete  
 DEFINITION  
 ACCESSION AL353685  
 VERSION AL353685.23 GI:14329534  
 KEYWORDS HTG.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 129608)  
 AUTHORS Tracey,A.  
 TITLE Direct Submission  
 JOURNAL Submitted (01-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk  
 requests: clonerequest@sanger.ac.uk  
 COMMENT  
 On Jun 8, 2001 this sequence version replaced gi:14272260.  
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at  
 http://www.sanger.ac.uk/Projects/C.elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping Group. Further information can be found at  
 http://www.sanger.ac.uk/HGP/Chr9  
 RP11-31J20 is from the library RPCI-11.1 constructed by the group

of Pieter de Jong. For further details see  
<http://www.chori.org/bacpac/home.htm>  
 VECTOR: pBACe3.6

IMPORTANT: this sequence is not the entire insert of clone  
 RP11-31J20 it may be shorter because we sequence overlapping  
 sections only once, except for a 100 base overlap.  
 The true right end of clone RP11-31J20 is at 129608 in this  
 sequence. The true right end of clone RP11-413C10 is at 2000 in  
 this sequence.

## FEATURES

## Location/Qualifiers

## Source

1. 129608  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="9"  
 /clone="RP11-31J20"  
 /clone\_lib="RPC1-11.1"

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

repeat\_region

/note="AluSq repeat: matches 1. .302 of consensus"  
 31424. .31734  
 /note="AluSp repeat: matches 1. .308 of consensus"  
 31987. .32116  
 /note="MER5A repeat: matches 60. .188 of consensus"  
 32450. .32536  
 /note="MER5A repeat: matches 14. .106 of consensus"  
 33293. .33380  
 /note="22 copies 4 mer atgg 79% conserved"  
 34435. .34588  
 /note="MER5A repeat: matches 13. .189 of consensus"  
 34729. .34873  
 /note="MIR repeat: matches 49. .198 of consensus"  
 35802. .35951  
 /note="MIR repeat: matches 77. .248 of consensus"  
 37183. .37260  
 /note="2 copies 39 mer 92% conserved"  
 37673. .37980  
 /note="AluSc repeat: matches 1. .308 of consensus"  
 39674. .40243  
 /note="L1MD2 repeat: matches 5774. .6331 of consensus"  
 40256. .40534  
 /note="L2 repeat: matches 2256. .2533 of consensus"  
 41476. .41615  
 /note="MIR repeat: matches 30. .185 of consensus"  
 42010. .42194  
 /note="MIR repeat: matches 49. .234 of consensus"  
 44809. .45239  
 /note="LTR16A repeat: matches 6. .450 of consensus"  
 45359. .45486  
 /note="MIR repeat: matches 1. .139 of consensus"  
 46654. .46693  
 /note="10 copies 4 mer acac 97% conserved"  
 47118. .47429  
 /note="AluSx repeat: matches 1. .312 of consensus"  
 47915. .48083  
 /note="AluSq repeat: matches 145. .313 of consensus"  
 49097. .49181  
 /note="L2 repeat: matches 2625. .2710 of consensus"  
 49578. .49758  
 /note="MIR repeat: matches 9. .188 of consensus"  
 49847. .49873  
 /note="MIR repeat: matches 155. .182 of consensus"  
 49874. .50188  
 /note="AluJb repeat: matches 3. .311 of consensus"  
 50189. .50347  
 /note="MIR repeat: matches 5. .155 of consensus"  
 51677. .51977  
 /note="AluSx repeat: matches 1. .303 of consensus"  
 52025. .52123  
 /note="L2 repeat: matches 1655. .1753 of consensus"  
 54337. .54544  
 /note="MIR repeat: matches 25. .246 of consensus"  
 54754. .55032  
 /note="AluSp repeat: matches 1. .302 of consensus"  
 55042. .55343  
 /note="AluV repeat: matches 5. .298 of consensus"  
 56027. .56232  
 /note="MIR repeat: matches 21. .259 of consensus"  
 56233. .56286  
 /note="18 copies 3 mer tgt 72% conserved"  
 56342. .56673  
 /note="L1MA9 repeat: matches 5893. .6307 of consensus"  
 56750. .56889  
 /note="L2 repeat: matches 2565. .2710 of consensus"  
 57406. .57567  
 /note="MER5A repeat: matches 4. .187 of consensus"  
 57864. .58174  
 /note="AluSq repeat: matches 1. .311 of consensus"  
 59443. .60004  
 /note="L2 repeat: matches 2088. .2723 of consensus"  
 60201. .60355  
 /note="MIR repeat: matches 20. .183 of consensus"



```

* 45549 48116: contig of 2568 bp in length
* 48117 48216: gap of 100 bp
* 52618: contig of 4402 bp in length
* 52719 52718: gap of 100 bp
* 56592: contig of 3874 bp in length
* 56993 56992: gap of 100 bp
* 56936 56935: contig of 2943 bp in length
* 59735 59734: gap of 100 bp
* 63661 63660: contig of 3926 bp in length
* 63662 63761: gap of 100 bp
* 63762 68437: contig of 4676 bp in length
* 68438 68537: gap of 100 bp
* 68538 71458: contig of 2921 bp in length
* 71459 71559: gap of 100 bp
* 71559 76888: contig of 5330 bp in length
* 76889 76988: gap of 100 bp
* 82113 82112: contig of 5125 bp in length
* 82214 82213: gap of 100 bp
* 82214 88220: contig of 6007 bp in length
* 88221 88320: gap of 100 bp
* 88321 93499: contig of 5179 bp in length
* 93500 93599: gap of 100 bp
* 93600 97901: contig of 4302 bp in length
* 97902 98001: gap of 100 bp
* 98002 103016: contig of 5015 bp in length
* 103017 103116: gap of 100 bp
* 103117 109178: contig of 6062 bp in length
* 109179 109278: gap of 100 bp
* 109279 117307: contig of 8029 bp in length
* 117308 117407: gap of 100 bp
* 117408 124079: contig of 6672 bp in length
* 124080 124179: gap of 100 bp
* 124180 131281: contig of 7102 bp in length
* 131282 131381: gap of 100 bp
* 131382 138059: contig of 6678 bp in length
* 138060 138159: gap of 100 bp
* 138160 145491: contig of 7332 bp in length
* 145492 145591: gap of 100 bp
* 145592 157391: contig of 11800 bp in length
* 157392 157491: gap of 100 bp
* 157492 175064: contig of 17573 bp in length.
FEATURES
    source
        1..175064 /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="RP11-1M10"
        /clone_lib="RP11-11 Human Male BAC"
        1..1003
            /note="assembly_fragment"
            1104..2634 /note="assembly_fragment"
            2735..4415 /note="assembly_fragment"
            4516..5785 /note="assembly_fragment"
            5886..7879 /note="assembly_fragment"
            7980..9686 /note="assembly_fragment"
            9787..12253 /note="assembly_fragment"
            12354..15228 /note="assembly_fragment"
            15329..17200 /note="assembly_fragment"
            17301..20131 /note="assembly_fragment"
            20232..22587 /note="assembly_fragment"
            22688..25707 /note="assembly_fragment"
            25808..28184 /note="assembly_fragment"

```

```

misc_feature 28285..31338 /note="assembly_fragment"
misc_feature 31439..34299 /note="assembly_fragment"
misc_feature 34400..38318 /note="assembly_fragment"
misc_feature 38419..42835 /note="assembly_fragment"
misc_feature 42936..45448 /note="assembly_fragment"
misc_feature 45549..48116 /note="assembly_fragment"
misc_feature 48217..52618 /note="assembly_fragment"
misc_feature 52719..56592 /note="assembly_fragment"
misc_feature 56693..59635 /note="assembly_fragment"
misc_feature 59736..63661 /note="assembly_fragment"
misc_feature 63762..68437 /note="assembly_fragment"
misc_feature 68538..71458 /note="assembly_fragment"
misc_feature 71559..76888 /note="assembly_fragment"

Query Match 95.3%; Score 340.2; DB 2; Length 175064;
Best Local Similarity 98.9%; Pred. No. 7.6e-93;
Matches 353; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 ttgagcttcagctagagagctgattagcagctcattgtatgtatgttcagca 60
Db 151511 TGAGGCTCAGCTAGAGGGCTGGATTAGCAGCTCCTATTGGTGTATGGCTTGACGA 151452

QY 61 ataatcgtatgctgtttccctccctcgtcttattcttcaatgaatgaccagccagcggt 120
Db 151451 ATAACTGATGGCTGTTTCCCTCCTGCTTTATCTTTCACTTAATGACCAAGCCAC -GGCGT 151393

QY 121 cctcgtcagctcgtgcccgtgcttcagggctcccgagccacacgctggcggtgt 180
Db 151392 CCCTGCTGTGAGCTGTGGCCGCTGCCCTTCAGGGGTCCCGAGCCACAGCCTGGGGTGTCT 151333

QY 181 gctcaggaacatgcatgttgctgagctgaggtgctgctgctgctgctgctgctgctcact 240
Db 151332 GGCTGAGGGAACATGGCTTGTGGCCCTCAGCTGAGTTCCTGCTGTGGGAAGAACCTCACT 151273

QY 241 ttcagaagaagacaaacagtaagcttgggtttttcagcagcgggggttctctcatctttt 300
Db 151272 TTCAGAAGAAGACAAACAGTAAGCTTGGGTTTTTCAGCAGCGGGGGTTCCTCATTTTTT 151213

QY 301 ttttgggttttggattgggattggagagggggagggagggaggaagagctgtgttg 357
Db 151212 TCTTTGTGTTTTTGATTGGGGATTGGAGGAGGGAGGGAGGAGGAGCTGTGTTG 151156

RESULT 7
AF287262 201144 bp DNA linear PRI 29-APR-2001
LOCUS Homo sapiens ATP-binding cassette 1 sub-family A member 1 (ABCA1)
DEFINITION and SNAP protein genes, complete cds.
ACCESSION AF287262
VERSION AF287262.1 GI:13876612
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 201144)
AUTHORS Qiu,Y., Cavellier,L., Chiu,S., Yang,X., Rubin,E. and Cheng,J.F.
TITLE Human and mouse abcal comparative sequencing and transgenesis
studies revealing novel regulatory sequences

```

JOURNAL MEDLINE REFERENCE AUTHORS TITLE JOURNAL	Genomics 73 (1), 66-76 (2001) 21251004 2 (bases 1 to 201144) Qiu,Y., Cavelier,L., Chiu,S., Rubin,E. and Cheng,J.-F. Direct Submission Submitted (13-JUL-2000) Genome Science Department, Lawrence Berkeley National Laboratory, 1 Cyclotron Rd, MS 84-171, Berkeley, CA 94720, USA
FEATURES	Location/Qualifiers
source	1..201144 /organism="Homo sapiens" /db_xref="taxon:9606" join(33931..34151,58319..58476,72893..72986,77520..77661, 78931..79049,101048..101169,104152..104328,117287..117379, 12329..12569,125285..125424,125754..125870, 130079..130276,131023..131228,131749..131925, 133708..133930,135677..135898,136959..137163, 138268..138381,140179..140350,141340..141471, 142777..142919,143123..143260,143963..144183, 145441..145513,146657..146859,148527..148575, 148771..148884,150280..150428,152078..152202, 153438..153536,156568..156757,158278..158372, 159682..159714,160837..160942,162417..162491, 163013..163182,164413..164590,166588..166703, 166815..166959,168505..168628,169715..169844, 170109..170229,171017..171079,171987..172093, 174448..174589,174961..175095,176039..176142, 176625..176717,177376..177619,178560..181457) /gene="ABCA1"
gene	/product="ATP-binding cassette 1 sub-family A member 1" 33931..181457 /gene="ABCA1"
exon	33931..34151 /gene="ABCA1"
exon	58319..58476 /gene="ABCA1"
CDS	join(58411..58476,72893..72986,77520..77661,78931..79049, 101048..101169,104152..104328,117287..117379, 12329..12569,125285..125424,125754..125870, 130079..130276,131023..131228,131749..131925, 133708..133930,135677..135898,136959..137163, 138268..138381,140179..140350,141340..141471, 142777..142919,143123..143260,143963..144183, 145441..145513,146657..146859,148527..148575, 148771..148884,150280..150428,152078..152202, 153438..153536,156568..156757,158278..158372, 159682..159714,160837..160942,162417..162491, 163013..163182,164413..164590,166588..166703, 166815..166959,168505..168628,169715..169844, 170109..170229,171017..171079,171987..172093, 174448..174589,174961..175095,176039..176142, 176625..176717,177376..177619,178560..178700) /gene="ABCA1" /codon_start=1 /product="ATP-binding cassette 1 sub-family A member 1" /protein_id="AAK43526.1" /db_xref="GI:13876613"
	/translation="MACWPOLLKLLKNTERRROTCLLLEAVPLFIFILISVRL SYPPYEQEHPKMPASCTLPWVQGLICNANNPCRYPTCEAGVGNFNAKSI ARLFSARDRLLYSKDSKMDRKLRLTLOIKSSNNKLQDFLVDFNEFSGFLYH KLSLPKSTVDKMLRADVLHKLHVFLQYQLHTSLCSKSEMIQLQDQVSELCGLP KELAAAEERLNNIDILKPIRLTNSIPFSEKELAEATKTLHSLGTLAQELFSMR SWSDMRQVMTNTVNSSSTQIYQAVSRVCGHPEGGGLKIKSLNWNEDNFKALF GGNGTEADAEFTVNDSTPYCNDMLNLESSPLRIWKALKPLLVGLYTPDTPAT ROVMAEVNKTFOELAVFHDLCGMEELSPKIWTPEMSEMDLVMLDLSRDHFEW QQLDGLDWTADQIVAFALAKHPEDVOSSNGSVYTRAEAFNETNQARTISRWCNVLN KLEPATEVNLINKSMELDERKFWAGIVFTGTPGSELPHVVKIKRMDINDVERT NKIKDQWDPGRDPEDMRYVWGGFAYLQDVQEQAIRVLTGTEKTKTYMQOQY PCYDIFLVRMSRNPFLMTLAWISYVAVIKGIYEKARLKEKTRIMRIGLDSILW FSWFTSLPLVSLAGLLVILKLGILLNPVSPDSVVFVLSFAVVTILQCFLLSTLF SRANLAAACGGIIFTLYLPPYVLCVWQDYVGFTLKFASLLSPVAFGFCCEFALFE
	EOGIGVQMDNLFSPVEEDGFNLTTSVSMMLFNFTLYGVMTWTVEAVFPQGVGIPRPW YFPCKTSYFGEESDEKSPGNSQKRISICEMEEPTHLKLGVSQNLVYVDMGMKV AVDGLANFYEGOITSFLHNGAGKTTMSILTGLPPTSGTAYILGKDIRSMKSTIR QNLGVCPOHNVLFDMLTVEHIFWYARLKLKSKKHYKAEMQMALDVGLPSSKLKST SOLSGMQRNLVALAFVGGSKVYILDEPTAGVDPYRRGICWELLKLYRQGTIIIST HHMDEADGDIRAIISHGKLCVGSLLKNOLGICGYTLTLVKKDYESSLSRNS STVSYLKDESDVSQSSDAGLSDHSDTLTIDVSAISNLIRKHVSEARLVEDIGHEL TYVLYPAAKEGAFVELFHEIDRLSLGTSISYGETTLEEFLLKVAEESGVDAETS DGLTPARNRRARFQKOSLPPTEDDADPNDSIDIPESRETDLLSGMDKGSYQVK GWLTKQOFVALLKRLLIARRSKGFAQIVLPVAVFCIALVFSIIVPPFGQYPSLE LQPMWNEQYTFVSNADPDTGLELLNALTKDPGEGTCMEGNPIPTPCQAGEEEM TTAPVPTIMDLFGNGNWTNPNPSPACQSSDKKMLKMLPCVPCGAGLPPQKQNTA DIQDITGRNISDYLKTYVQIIAKSLNNKWNEFRIYGFSLGVNTQALPPSQEVN DAILQKKHLKLAKSDSADFLNSLGRMTGLDTRNNKVKFNKMGHAISSIPAMFV NAILRNLQKGNPSHYGITAFNHNPLNLTQOLSEVALMTTSVDVLVSVICVIFAMSFV PASVFNLIQERYSKAKHLQFISGVKPIVWLSFVDMCNVYVPATLVIIIFIGQV KSVYSTNLPVLAIIIIIGWTSITPLMPASPFVKIPSTAYVVLTVNLIIFIGNSVA TFVLELFTDNKNINDILKASFLIPFPHFCGLRGLDMKVNQMAADALERFGENRVS PLSDWLVGRNLFAMAVEGVVFFLITVLIQTRFIRPRPVNAKLSPLNDEDEDVRRKQ RILDGGQNDILEIKELTKIYRRKRPADVRCICVPPGECFGLLGVNGAGKSTFKM LTGDTTVTRGDALFNKNSILSNIEHVQNNGYCQDFDAITELTGREHVEFFALLRGV PEKEVGVGEWAIKRLGLVKYGEKYNKSGNKRKLSLTAALIGGPPVVFIDEPPTG MDPKARFLNLCALSYYVKEGRSVLTSMSMECEALCTMAIWNVGRFCGLGVOHLK NRFCDGTYIVRIAGSNPDLKPVQDFGLAPPGSVLKEKHRNMLQYOLPSSLSLARI FSILSQSKRLHIEDYSVSQTTLDQVFNFAKQSDDDHLKLSLHKNQITVDVAVLT SFLQDEKVKESYY" 72893..72986 /gene="ABCA1" /number=3 77520..77661 /gene="ABCA1" /number=4 78931..79049 /gene="ABCA1" /number=5 101048..101169 /gene="ABCA1" /number=6 104152..104328 /gene="ABCA1" /number=7 117287..117379 /gene="ABCA1" /number=8 12329..12569 /gene="ABCA1" /number=9 125285..125424 /gene="ABCA1" /number=10 125754..125870 /gene="ABCA1" /number=11 130079..130276 /gene="ABCA1" /number=12 131023..131228 /gene="ABCA1" /number=13 131749..131925 /gene="ABCA1" /number=14 133708..133930 /gene="ABCA1" /number=15 135677..135898 /gene="ABCA1" /number=16 136959..137163 /gene="ABCA1" /number=17 138268..138381 /gene="ABCA1"

exon	/number=18 140179. .140350 /gene="ABCA1" /number=19 141340. .141471 /gene="ABCA1" /number=20 142777. .142919 /gene="ABCA1" /number=21 143123. .143260 /gene="ABCA1" /number=22 143963. .144183 /gene="ABCA1" /number=23 145441. .145513 /gene="ABCA1" /number=24 146657. .146859 /gene="ABCA1" /number=25 148527. .148575 /gene="ABCA1" /number=26 148771. .148884 /gene="ABCA1" /number=27 150280. .150428 /gene="ABCA1" /number=28 152078. .152202 /gene="ABCA1" /number=29 153438. .153536 /gene="ABCA1" /number=30 156568. .156757 /gene="ABCA1" /number=31	Query Match 95.3%; Score 340.2; DB 9; Length 201144; Best Local Similarity 98.9%; Pred. No. 7.7e-93; Matches 353; Conservative 0; Mismatches 3; Indels 1; Gaps 1;	
exon	Qy 1 tgagagctcagctgagagggctgattagcagctcctcatgtgtgtatggctttgcacga 60       Db 58220 TGAGAGGCTCAGCTGAGAGGGCTGGGATTAAGCAGTCTCTAATGGTGTATGGCTTTGCACGA 58279 		
exon	Qy 61 ataactgatgctgtttccctcctctctcttattcttcaagtaataaccagcagcagggcggt 120       Db 58280 ATAACGATGCGCTTTCCCTCCCTCTTATCTTTCAGTAAATCACCAGCCAC-GGCGT 58338 		
exon	Qy 121 cctgtgtcagctctgcccgtctgcttccaggggtcccgagccacacgctggcggtgtct 180       Db 58339 CCCTGTGTGAGCTCTGGCGGTGCTCTCCAGGGCTCCGAGCCACAGCGTGGGGGTGCT 58398 		
exon	Qy 181 gqctgaggaacatgcattgttggcctcagctcaggttgcctctgtctgtgaagaacctact 240       Db 58399 GGCTGAGGGAACATGCTGTGTGGCCTCAGCTGAGGTTGCTGCTGTGGAAGAACCTCACT 58458 		
exon	Qy 241 ttcagaagaagacaaacagtaagcttgggtttttcagcagcgggggttctctcaattttt 300       Db 58459 TTCAGAGAAGACAAACAGTAAGCTTGGGTTTTTCAGCAGCGGGGGTCTCTCATTTT 58518 		
exon	Qy 301 tctttgtgttttgaattgggattggagggaggagggaagaaagctgtgttg 357       Db 58519 TCTTTGTGTTTTCAGTTGGGATTTGGAGGGAGGGAGGGAAGGAAAGCTGTGTG 58575 		
RESULT 8	AC021345/c	90698 bp	DNA
AC021345/c	LOCUS	linear	HTG 13-JUL-2000

44357	44456:	gap of	100 bp
44457	45325:	contig of 869 bp	in length
45326	45425:	gap of	100 bp
45426	46305:	contig of 880 bp	in length
46306	46405:	gap of	100 bp
46406	47302:	contig of 897 bp	in length
47303	47402:	gap of	100 bp
47403	48293:	contig of 891 bp	in length
48294	48393:	gap of	100 bp
48394	49280:	contig of 887 bp	in length
49281	49380:	gap of	100 bp
49381	50250:	contig of 870 bp	in length
50251	50350:	gap of	100 bp
50351	51209:	contig of 859 bp	in length
51210	51309:	gap of	100 bp
51310	52221:	contig of 912 bp	in length
52222	52321:	gap of	100 bp
52322	53205:	contig of 884 bp	in length
53206	53305:	gap of	100 bp
53306	54225:	contig of 920 bp	in length
54226	54325:	gap of	100 bp
54326	55206:	contig of 881 bp	in length
55207	55306:	gap of	100 bp
55307	56165:	contig of 859 bp	in length
56166	56285:	gap of	100 bp
56286	57124:	contig of 859 bp	in length
57125	57224:	gap of	100 bp
57225	58121:	contig of 897 bp	in length
58122	58221:	gap of	100 bp
58222	59075:	contig of 854 bp	in length
59076	59175:	gap of	100 bp
59176	60058:	contig of 883 bp	in length
60059	60158:	gap of	100 bp
60159	61067:	contig of 909 bp	in length
61068	61167:	gap of	100 bp
61168	62051:	contig of 884 bp	in length
62052	62151:	gap of	100 bp
62152	63022:	contig of 871 bp	in length
63023	63122:	gap of	100 bp
63123	64026:	contig of 904 bp	in length
64027	64126:	gap of	100 bp
64127	65015:	contig of 889 bp	in length
65016	65115:	gap of	100 bp
65116	65975:	contig of 860 bp	in length
65976	66075:	gap of	100 bp
66077	66954:	contig of 879 bp	in length
66955	67054:	gap of	100 bp
67055	67947:	contig of 893 bp	in length
67948	68047:	gap of	100 bp
68048	68934:	contig of 887 bp	in length
68935	69034:	gap of	100 bp
69035	69910:	contig of 876 bp	in length
69911	70010:	gap of	100 bp

3

60 13 12 13 18 12 24 12



QY 241 ttcaagaagacacaaagtaagctgtgtgtttttc-agcagcggggtttctc 292  
 |||||  
 Db 12931 TTCAGAGAGACAAACAGTAGCTTGGTGTTCATGACGACGGGGGTTC 12879  
 |||||  
 RESULT 9  
 AF258623S2  
 LOCUS  
 DEFINITION Homo sapiens ATP binding cassette transporter 1 (ABCA1) gene, exon 2.  
 AF258623S2 200 bp DNA linear PRI 23-JUN-2000  
 VERSION AF258624.1 GI:7769714  
 KEYWORDS  
 SEGMENT 2 of 4  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 Pullinger, C.R., Hakamata, H., Duchateau, P.N., Eng, C.,  
 Aouizerat, B.E., Fielding, C.J., and Kane, J.P.  
 Analysis of hABC1 gene 5' end: additional peptide sequence,  
 promoter region, and four polymorphisms  
 Biochem. Biophys. Res. Commun. 271 (2000) In press  
 REFERENCE 2 (bases 1 to 200)  
 JOURNAL  
 AUTHORS Pullinger, C.R., Hakamata, H., Duchateau, P.N., Eng, C.,  
 Aouizerat, B.E., Fielding, C.J., and Kane, J.P.  
 TITLE Direct Submission  
 JOURNAL Submitted (23-JUN-2000) Cardiovascular Research Institute,  
 University of California, San Francisco, 505 Parnassus Avenue, San  
 Francisco, CA 94143-0130, USA  
 REMARK Sequence update by submitter  
 FEATURES  
 source  
 1..200  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="9"  
 /map="9q31"  
 22..179  
 /gene="ABCA1"  
 /number=2  
 exon  
 37 a 56 c 57 g 50 t  
 BASE COUNT 37 a 56 c 57 g 50 t  
 ORIGIN  
 Query Match 51.6%; Score 184.2; DB 9; Length 200;  
 Best Local Similarity 98.0%; Pred. No. 1.6e-45;  
 Matches 197; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
 QY 79 cccctcgtttatcttcaagtaataaccagccacggcgctccctgtcgtcagctctgg 138  
 |||||  
 Db 1 CCCCTCCTGCTTATCTTCAGTTAATGACACGCCAC-GGCGTCCCTGCTGTGAGCTCTGG 59  
 |||||  
 QY 139 ccgctgcctccaggggtcccgagccacacgctggggcgctgctgctgaggaacatggca 198  
 |||||  
 Db 60 CCGCTGCCTTCCAGGGCTCCCGACCCACACGCTGGGGGTGCTGGCTGAGGGAACATGGCT 119  
 |||||  
 QY 199 tgttgccctcagctgaggtgtcgtctgtgaagaacatcatttcagaagaacaaca 258  
 |||||  
 Db 120 TGTGGGCTCAGCTGAGGTGCTGCTGTGGGAAGAACCTCATTTCAGAAGAACAACA 179  
 |||||  
 QY 259 gtaagcttggtttttcagca 279  
 |||||  
 Db 180 GTAAGCTTGGGTTTTTCAGCA 200  
 |||||  
 RESULT 10  
 AX060713  
 LOCUS  
 DEFINITION Sequence 1 from Patent WO0078972.  
 AX060713 10442 bp DNA linear PAT 22-JAN-2001  
 VERSION AX060713.1 GI:12406103

## KEYWORDS

human.

## SOURCE

Homo sapiens

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 10442)

## AUTHORS

Lawn, R.M., Wade, D., and Garvin, M.

## TITLE

Regulation with binding cassette transporter protein abcl

## JOURNAL

Patent: WO 0078972-A 1 28-DEC-2000;

## FEATURES

CV THERAPEUTICS, INC. (US)

## source

Location/Qualifiers  
1..10442  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

## BASE COUNT

2898 a 2297 c 2408 g 2835 t 4 others

## ORIGIN

Query Match 44.6%; Score 159.4; DB 6; Length 10442;  
Best Local Similarity 81.8%; Pred. No. 1.2e-37;  
Matches 184; Conservative 0; Mismatches 41; Indels 0; Gaps 0;QY 97 cagttaatgaccagccacggcgctccctgtcagctctggcgctgctccagggct 156  
|||||Db 195 CAGTTAATGACCAACGACGGCGCTCCCTGCTGTGAGCTCTGGCGCTGCTTCCAGGGCT 254  
|||||QY 157 ccgagccacacgctggcgctgctgctgaggaacatggcatgttgccctcagctgaggg 216  
|||||Db 255 CCGGAGCCACACCTGGCGTGTGGCTGAGGGAACATGGCTTGTGGCTCAGCTCAGG 314  
|||||QY 217 ttgctgtgtggaacacctcacttcacagaagaacacagtaagcttggttttca 276  
|||||Db 315 TTGCTGCTGTGGAGAACCTCATTTCAGAAGAACAACATGTCAGCTGTGTACTGGAA 374  
|||||QY 277 gcagcggggggttctctcatctttttttgtgtgtttgtgagttggg 321  
|||||Db 375 GTGGCTGGCGCTCTATTATCTTCCTGATCCTGATCTCTGTCGG 419  
|||||

## RESULT 11

## AX060892

## LOCUS

Sequence 1 from Patent WO0078971.

## DEFINITION

AX060892

## ACCESSION

AX060892.1 GI:12406270

## VERSION

## KEYWORDS

human.

## SOURCE

Homo sapiens

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 10442)

## AUTHORS

Lawn, R.M., Wade, D., Oram, J.F. and Garvin, M.

## TITLE

Atp binding cassette transporter protein abcl polypeptides

## JOURNAL

Patent: WO 0078971-A 1 28-DEC-2000;

## FEATURES

CV THERAPEUTICS, INC. (US)

## source

Location/Qualifiers  
1..10442  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

## BASE COUNT

2898 a 2297 c 2408 g 2835 t 4 others

## ORIGIN

Query Match 44.6%; Score 159.4; DB 6; Length 10442;  
Best Local Similarity 81.8%; Pred. No. 1.2e-37;  
Matches 184; Conservative 0; Mismatches 41; Indels 0; Gaps 0;QY 97 cagttaatgaccagccacggcgctccctgtcagctctggcgctgctccagggct 156  
|||||Db 195 CAGTTAATGACCAACGACGGCGCTCCCTGCTGTGAGCTCTGGCGCTGCTTCCAGGGCT 254  
|||||QY 157 ccgagccacacgctggcgctgctgctgaggaacatggcatgttgccctcagctgaggg 216  
|||||







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 03:08:22 ; Search time 520.94 Seconds  
(without alignments)  
1176.601 Million cell updates/sec

Title: US-09-846-456-2  
Perfect score: 357  
Sequence: 1 tggaggtctcagctgaggg.....gaggaaggaagctgtgttg 357

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_032802.\*

- 1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1980.DAT.\*
- 2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1981.DAT.\*
- 3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1982.DAT.\*
- 4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1983.DAT.\*
- 5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1984.DAT.\*
- 6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1985.DAT.\*
- 7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1986.DAT.\*
- 8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1987.DAT.\*
- 9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1988.DAT.\*
- 10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1989.DAT.\*
- 11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1990.DAT.\*
- 12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1991.DAT.\*
- 13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1992.DAT.\*
- 14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1993.DAT.\*
- 15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1994.DAT.\*
- 16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1995.DAT.\*
- 17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1996.DAT.\*
- 18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1997.DAT.\*
- 19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1998.DAT.\*
- 20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1999.DAT.\*
- 21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA2000.DAT.\*
- 22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT.\*
- 23: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA2001B.DAT.\*
- 24: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	351.2	98.4	183999	22	AAF92831 Human ABC1 genomic
2	340.2	95.3	10545	21	AAC69132 Human ABC1 gene ex
3	159.4	44.6	10442	22	AAF24680 Nucleotide sequenc
4	159.4	44.6	10442	22	AAF24702 Nucleotide sequenc
5	159.4	44.6	10474	22	AAF24685 Nucleotide sequenc
6	159.4	44.6	10474	22	AAF24686 Nucleotide sequenc
7	159.4	44.6	10474	22	AAF24707 Nucleotide sequenc
8	159.4	44.6	10474	22	AAF24708 Nucleotide sequenc
9	145.8	40.8	446	22	AA504035 Partial human ABC1

10	145.8	40.8	7086	22	ABA09200 Human ABCA1 homolo
11	145.8	40.8	7086	22	AAK52667 Human polynucleoti
12	145.8	40.8	7260	22	AAJ21326 Human ATP binding
13	145.8	40.8	7260	22	AAJ70315 Human polynucleoti
14	145.8	40.8	7260	22	AAK51683 Human polynucleoti
15	145.8	40.8	9741	22	AAS06120 Human ABC1 DNA seq
16	145.8	40.8	9854	22	AAS06120 Human ABC1 DNA seq
17	144.2	40.4	736	22	AAH07432 Human ABC1 CDNA clone
18	144.2	40.4	1556	22	AAH18606 Human CDNA sequenc
19	135.8	38.0	7857	21	AAC69388 Human ABC1 cholest
20	135.8	38.0	7860	22	AAF83826 Human ABC1 cholect
21	135.8	38.0	7860	22	AAF83826 Human ABC1 cholect
22	135.8	38.0	7861	21	AAF83826 Human ABC1 cholect
23	135.8	38.0	7861	21	AAF83826 Human ABC1 cholect
24	135.8	38.0	7864	21	AAC69385 Human ABC1 cholect
25	135.8	38.0	7864	21	AAC69385 Human ABC1 cholect
26	135.8	38.0	7864	21	AAC69385 Human ABC1 cholect
27	42.6	11.9	23024	22	AAF25499 Nucleotide sequenc
28	36.4	10.2	38	22	AAF93082 ABC1 polymorphism
29	36	10.1	7661	22	AAJ16545 Human novel protei
30	36	10.1	7661	22	AAJ16545 Human genomic DNA
31	36	10.1	7661	22	AAJ16545 Human polynucleoti
32	36	10.1	11696	22	AAJ16546 Human novel protei
33	36	10.1	11696	22	AAJ16546 Human novel protei
34	36	10.1	11696	22	AAJ16546 Human novel protei
35	36	10.1	11696	22	AAJ16546 Human novel protei
36	36	10.1	11696	22	AAJ16546 Human novel protei
37	36	10.1	11696	22	AAJ16546 Human novel protei
38	36	10.1	11696	22	AAJ16546 Human novel protei
39	36	10.1	11696	22	AAJ16546 Human novel protei
40	36	10.1	11696	22	AAJ16546 Human novel protei
41	36	10.1	11696	22	AAJ16546 Human novel protei
42	35.8	10.0	534	21	AAJ75311 Human ORFX ORF866
43	35.8	10.0	1133	22	AAJ75311 Human secreted pro
44	35.8	10.0	50885	22	AAJ75311 Human immune/haema
45	35.8	10.0	92407	22	AAJ75311 Genomic fragment #

# ALIGNMENTS

RESULT 1

AAF92831

ID AAF92831 standard; DNA; 183999 BP.

XX AAF92831;

XX AAF92831;

DT 17-MAY-2001 (first entry)

XX Human ABC1 genomic DNA.

DE High density lipoprotein-cholesterol; HDL-C; cardiovascular; ABC1; ds.

XX Homo sapiens.

XX WO200115676-A2.

XX 08-MAR-2001.

XX 01-SEP-2000; 2000WO-IB01492.

XX 01-SEP-1999; 99US-0151977.

XX 15-MAR-2000; 2000US-0526193.

XX 23-JUN-2000; 2000US-0213958.

XX (UYBR-) UNIV BRITISH COLUMBIA.

XX (XENO-) XENON GENETICS INC.

XX Hayden MR, Brooks-Wilson AR, Pimstone SN, Clee SM;

XX WPI; 2001-244356/25.

XX Treating a lower than normal high density lipoprotein-cholesterol

PT



Db 8162 cctgtgtgagctctgcccgtcttccagggctcccgagccacacagcgtgggtgct 8221  
 QY 181 ggctagaggaacatgcagctgttgctcagctcagctgaggtgtgctgtggaagaacctcaact 240  
 Db 8222 ggcctagaggaacatgcagctgttgctcagctcagctgaggtgtgctgtggaagaacctcaact 8281  
 QY 241 ttccagaagaagaacacagtaagcttggtttttcagcagcgggggttctctcattttt 300  
 Db 8282 ttccagaagaagaacacagtaagcttggtttttcagcagcgggggttctctcattttt 8341  
 QY 301 tcttctggtttgttggttggttggttggttggttggttggttggttggttggttggttggt 357  
 Db 8342 tcttctggtttgttggttggttggttggttggttggttggttggttggttggttggttggt 8398

RESULT 3  
 AAF24680  
 ID AAF24680 standard; DNA; 10442 BP.

XX AC AAF24680;

DT 20-APR-2001 (first entry)

XX Nucleotide sequence of a human ABC1 polypeptide.

XX Human; adenosine triphosphate binding cassette protein 1; ABC1;  
 KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;  
 KW chromosome 9q22-9q31; heart disease; hypercholesterolemia;  
 KW atherosclerosis; cholesterol transport; ss.

OS Homo sapiens.

XX Key Location/Qualifiers  
 FH CDS 291..7076

FT /\*tag= a  
 FT /product= "ABC1 polypeptide"

XX WO200078972-A2.

XX 28-DEC-2000.

XX 16-JUN-2000; 2000WO-US16765.

XX 18-JUN-1999; 99US-0140264.

XX 14-SEP-1999; 99US-0153872.

XX 19-NOV-1999; 99US-0166573.

XX (CVTH-) CV THERAPEUTICS INC.

XX Lawn RM, Wade D, Garvin M;

XX WPI; 2001-137812/14.

XX Adenosine triphosphate (ATP) binding cassette (ABC) polynucleotide,  
 PT useful for the development of agents for the treatment of heart disease  
 PT and other disorders associated with hypercholesterolemia and  
 PT atherosclerosis -

PS Disclosure; Page 122-128; 215pp; English.

XX The present sequence encodes a human adenosine triphosphate (ATP)  
 CC binding cassette protein (ABC) 1 polypeptide. ABC1 resides in cell  
 CC membranes and utilises ATP hydrolysis to transport a wide variety of  
 CC substrates across the plasma membrane. ABC1 is a pivotal protein in  
 CC the apolipoprotein-mediated mobilisation of intracellular cholesterol  
 CC stores. ABC1 is defective in Tangier disease, a genetic disorder  
 CC characterised by abnormal HDL-cholesterol metabolism. The ABC1 gene is  
 CC localised to chromosome 9q22-9q31. The ABC1 genes and proteins are  
 CC useful for developing pharmaceutical agents for the treatment of heart  
 CC disease and other disorders associated with hypercholesterolemia and  
 CC atherosclerosis. The genes are useful for developing screening assays to  
 CC screen for compounds that regulate the expression of genes associated  
 CC with cholesterol transport. The genes and proteins are also useful for

CC are also useful as diagnostic indicators of cardiovascular disease and  
 CC other disorders associated with hypercholesterolemia.  
 XX Sequence 10442 BP; 2898 A; 2297 C; 2408 G; 2835 T; 4 other;

Query Match 44.6%; Score 159.4; DB 22; Length 10442;  
 Best Local Similarity 81.8%; Pred. No. 9.3e-39;  
 Matches 184; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 97 cagtaataaccagccacggcgctccctgtgtcagctgtgcgcgtccctccaggct 156  
 Db 195 cagtaataaccagccacggcgctccctgtgtcagctgtgcgcgtccctccaggct 254  
 QY 157 ccgagaccacacgctggcgctgtgctgagggaacatggcatgttgctcagctgagg 216  
 Db 255 ccgagaccacacgctggcgctgtgctgagggaacatggcatgttgctcagctgagg 314  
 QY 217 ttgctgtgtggaagaacctcactttcagaagaagacacacagtaagcttgggttttca 276  
 Db 315 ttgctgtgtggaagaacctcactttcagaagaagacacacagtaagcttgggttttca 374  
 QY 277 gcagcgggggttctctcattttttttttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 321  
 Db 375 gtggcgtgctctatttattcttct 419

RESULT 4

AAF24702

ID AAF24702 standard; DNA; 10442 BP.

XX AC AAF24702;

DT 20-APR-2001 (first entry)

XX Nucleotide sequence of a human ABC1 polypeptide.

XX Human; adenosine triphosphate binding cassette protein 1; ABC1;  
 KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;  
 KW chromosome 9q22-9q31; heart disease; hypercholesterolemia;  
 KW atherosclerosis; cholesterol transport; ss.

OS Homo sapiens.

XX Key Location/Qualifiers  
 FH CDS 291..7076

FT /\*tag= a  
 FT /product= "ABC1 polypeptide"

XX WO200078971-A2.

XX 28-DEC-2000.

XX 16-JUN-2000; 2000WO-US16591.

XX 18-JUN-1999; 99US-0140264.

XX 14-SEP-1999; 99US-0153872.

XX 19-NOV-1999; 99US-0166573.

XX (CVTH-) CV THERAPEUTICS INC.

XX (UNIW) UNIV WASHINGTON.

XX Lawn RM, Wade D, Oram JF, Garvin M;

XX WPI; 2001-137811/14.

XX P-PSDB; AAB31365.

XX Adenosine triphosphate (ATP) binding cassette protein (ABC) 1  
 PT polynucleotides and polypeptides, useful for treatment of heart disease  
 PT and other disorders associated with hypercholesterolemia and  
 PT atherosclerosis -  
 XX Claim 3; Page 117-123; 211pp; English.





```

FT FT      /*tag= a
XX FT      /product= "defective ABC1 polypeptide"
XX PN      WO200078972-A2.
XX PD      28-DEC-2000.
XX PF      16-JUN-2000; 2000WO-US16765.
XX PR      18-JUN-1999; 99US-0140264.
XX PR      14-SEP-1999; 99US-0153872.
XX PR      19-NOV-1999; 99US-0166573.
XX PA      (CVTH-) CV THERAPEUTICS INC.
XX PI      Lawn RM, Wade D, Garvin M;
XX PN      WPI; 2001-137812/14.
XX DR      Adenosine triphosphate (ATP) binding cassette (ABC) polynucleotide,
XX PT      useful for the development of agents for the treatment of heart disease
XX PT      and other disorders associated with hypercholesterolemia and
XX PT      atherosclerosis -
XX PS      Disclosure; Page 170-176; 215pp; English.
XX CC      The present sequence encodes a human adenosine triphosphate (ATP)
XX CC      binding cassette protein (ABC) 1 polypeptide, and is isolated from
XX CC      a Tangier disease patient. ABC1 resides in cell membranes and utilises
XX CC      ATP hydrolysis to transport a wide variety of substrates across the
XX CC      plasma membrane. ABC1 is a pivotal protein in the apolipoprotein-mediated
XX CC      mobilisation of intracellular cholesterol stores. ABC1 is defective in
XX CC      Tangier disease, a genetic disorder characterised by abnormal
XX CC      HDL-cholesterol metabolism. The ABC1 gene is localised to chromosome
XX CC      9q22-q31. The ABC1 genes and proteins are useful for developing
XX CC      pharmaceutical agents for the treatment of heart disease and other
XX CC      disorders associated with hypercholesterolemia and atherosclerosis. The
XX CC      genes are useful for developing screening assays to screen for compounds
XX CC      that regulate the expression of genes associated with cholesterol
XX CC      transport. The genes and proteins are also useful for are also useful
XX CC      as diagnostic indicators of cardiovascular disease and other disorders
XX CC      associated with hypercholesterolemia.
XX SQ      Sequence 10474 BP; 2907 A; 2304 C; 2415 G; 2844 T; 4 other;

Query Match      44.6%; Score 159.4; DB 22; Length 10474;
Best Local Similarity 81.8%; Pred. No. 9.3e-39;
Matches 184; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 97 cagttaatgaccagcagcggtccctgctgtcagctgtggtccgctgctccagggt 156
   |||||
Db 227 cagttaatgaccagcagcggtccctgctgtgagctgtggtccgctgctccagggt 286
   |||||

QY 157 cccagaccacagctggcggtgctggtgaggaacatggcatgttggcctcagctgagg 216
   |||||
Db 287 cccagaccacagctggcggtgctggtgaggaacatggctgttggcctcagctgagg 346
   |||||

QY 217 ttgctgctgtggaagaacctcactttcagagaagacaacagtaagcttgggttttca 276
   |||||
Db 347 ttgctgctgtggaagaacctcactttcagagaagacaacatgctgcttactggaa 406
   |||||

QY 277 gcagcgggggtgtctctcattttttttttttgtgtgtttgagttggg 321
   |||||
Db 407 gtggcctggcctatttattcttcctgatcctgatctctgttctgg 451
   |||||

RESULT 7
AAF24707
ID AAF24707 standard; DNA; 10474 BP.
XX AC
XX AAF24707;
XX

```

```

DT 20-APR-2001 (first entry)
XX Nucleotide sequence of ABC1 polypeptide from Tangier disease patient.
DE
XX
XX Human; adenosine triphosphate binding cassette protein 1; ABC1;
KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
KW chromosome 9q22-q31; heart disease; hypercholesterolemia;
KW atherosclerosis; cholesterol transport; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 323..7108
XX FT /*tag= a
XX FT /product= "defective ABC1 polypeptide"
XX
XX WO200078971-A2.
XX
XX 28-DEC-2000.
XX
XX 16-JUN-2000; 2000WO-US16591.
XX
XX 18-JUN-1999; 99US-0140264.
XX 14-SEP-1999; 99US-0153872.
XX 19-NOV-1999; 99US-0166573.
XX
XX (CVTH-) CV THERAPEUTICS INC.
XX PA (UNIW ) UNIV WASHINGTON.
XX
XX Lawn RM, Wade D, Oram JF, Garvin M;
XX
XX WPI; 2001-137811/14.
XX P-PSDB; AAB31366.
XX
XX Adenosine triphosphate (ATP) binding cassette protein (ABC) 1
XX polynucleotides and polypeptides, useful for treatment of heart disease
XX and other disorders associated with hypercholesterolemia and
XX atherosclerosis -
XX
XX Claim 27; Page 144-150; 211pp; English.
XX
XX The present sequence encodes a human adenosine triphosphate (ATP)
XX binding cassette protein (ABC) 1 polypeptide, and is isolated from
XX a Tangier disease patient. ABC1 resides in cell membranes and utilises
XX ATP hydrolysis to transport a wide variety of substrates across the
XX plasma membrane. ABC1 is a pivotal protein in the apolipoprotein-mediated
XX mobilisation of intracellular cholesterol stores. ABC1 is defective in
XX HDL-cholesterol metabolism. The ABC1 gene is localised to chromosome
XX 9q22-q31. The ABC1 genes and proteins are useful for developing
XX pharmaceutical agents for the treatment of heart disease and other
XX disorders associated with hypercholesterolemia and atherosclerosis. The
XX genes are useful for developing screening assays to screen for compounds
XX that regulate the expression of genes associated with cholesterol
XX transport. The genes and proteins are also useful for are also useful
XX as diagnostic indicators of cardiovascular disease and other disorders
XX associated with hypercholesterolemia.
XX
XX Sequence 10474 BP; 2906 A; 2305 C; 2416 G; 2843 T; 4 other;

Query Match      44.6%; Score 159.4; DB 22; Length 10474;
Best Local Similarity 81.8%; Pred. No. 9.3e-39;
Matches 184; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 97 cagttaatgaccagcagcggtccctgctgtcagctgtggtccgctgctccagggt 156
   |||||
Db 227 cagttaatgaccagcagcggtccctgctgtgagctgtggtccgctgctccagggt 286
   |||||

QY 157 cccagaccacagctggcggtgctggtgaggaacatggcatgttggcctcagctgagg 216
   |||||
Db 287 cccagaccacagctggcggtgctggtgaggaacatggctgttggcctcagctgagg 346
   |||||

```







QY 157 cccgagccacacgctggcgctgctgctgaggaacatggaatgttgctcgaagctgagg 216  
 Db 285 cccgagccacacgctggcgctgctgaggaacatggaatgttgctcgaagctgagg 344  
 QY 217 ttgctgctgtggaagaacctcactttcagaagaagacacagtaagcttggttttca 276  
 Db 345 ttgctgctgtggaagaacctcactttcagaagaagacacagtaagcttggttttca 404  
 QY 277 gcagcgggggttctctcattttttttgtgtgtttgagttggg 321  
 Db 405 gtggcctggcctctatttctcctgacccgctctgttcgg 449

## RESULT 13

ID AAI70315 standard; cDNA; 7260 BP.

AC AAI70315;

DT 07-JAN-2002 (first entry)

XX Human ATP binding cassette transporter 1 (ABCI) cDNA.

XX ATP binding cassette transporter 1; ABC1; human; lipid disorder;  
 KW cholesterol; cardiovascular disease; inflammatory disease;  
 KW antiinflammatory; antilipemic; antipsoriatic; dermatological;  
 KW Tangier disease; coronary heart disease; diagnosis; gene therapy;  
 KW polymorphism; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 321..7106

FT /\*tag= a

FT CDS 501..7106

FT /\*tag= b

FT /\*note= "alternative open reading frame of AAI70314"

FT variation replace(976,A)

FT /\*tag= c

FT variation replace(1516,C)

FT /\*tag= d

FT variation replace(2969,G)

FT /\*tag= e

FT variation replace(3836,C)

FT /\*tag= f

FT EPI136554-A1.

PN 26-SEP-2001.

XX 24-MAR-2000; 2000EP-0106401.

XX 24-MAR-2000; 2000EP-0106401.

XX (FARB ) BAYER AG.

XX Schmitz G, Bodzioch M;

XX WPI; 2001-640389/74.

XX P-PSDB; RAM50228.

XX New adenosine triphosphate binding cassette transporter gene

XX polymorphisms, useful for diagnosing and treating lipid disorders,

XX cardiovascular diseases and inflammatory diseases

XX Disclosure; Page 26-28; 41pp; English.

XX The present sequence is that of cDNA encoding the human adenosine

XX triphosphate (ATP) binding cassette transporter 1 (ABCI) protein

XX (see RAM50227). The sequence includes an extended open reading

XX frame (ORF) to that provided by the sequence in AAI70314, using

XX an alternative ATG codon as initiation codon and thereby adding an

CC extra 40 N-terminal amino acids to the encoded ABC1 protein (see  
 CC RAM50228). The invention provides 4 common polymorphisms in the  
 CC ABC1 gene. These were identified by sequencing the ABC1 gene in  
 CC different Tangier kindreds. In the variant genes (numbering as in  
 CC AAI70314), G is changed to A at position 596, T is changed to C at  
 CC position 1136, A is changed to G at position 2589 or G is changed  
 CC to C at position 3456, or any combination of these. All of these  
 CC polymorphisms alter the amino acid sequence of ABC1 and therefore  
 CC may affect its function. The 2 most common polymorphisms (G596A)  
 CC and A2589G) are both associated with a decreased in vitro ApoA-I  
 CC mediated efflux of cholesterol from mononuclear phagocytes, a  
 CC feature typical of Tangier disease. 3 Of the variants (G596A,  
 CC A2589G and G3456C) are significantly increased in a population of  
 CC men having low high density lipoprotein-cholesterol levels and  
 CC established coronary heart disease (CHD) relative to CHD-free  
 CC control subjects. The use of the provided ABC1 polymorphisms for  
 CC the diagnosis and treatment of lipid disorders, cardiovascular  
 CC diseases, and inflammatory diseases (e.g. psoriasis, lupus  
 CC erythematoses) is claimed. Modulation of ABC1 transcripts or  
 CC proteins by antisense or ribozyme technology or RNA decoys is also  
 CC claimed.

XX Sequence 7260 BP; 1834 A; 1765 C; 1905 G; 1756 T; 0 other;

Query Match 40.8%; Score 145.8; DB 22; Length 7260;  
 Best Local Similarity 80.9%; Pred. No. 1.3e-34;

Matches 182; Conservative 0; Mismatches 42; Indels 1; Gaps 1;

QY 97 cagttaatgaccagccagcggtccctgctgctgagctggtgctgctcagggact 156

Db 226 cagttaatgaccagccagcggtccctgctgctgagctggtgctgctcagggact 284

QY 157 cccgagccacacgctggcgctgctgctgaggaacatggaatgttgctcagctgagg 216

Db 285 cccgagccacacgctggcgctgctgctgaggaacatggaatgttgctcagctgagg 344

QY 217 ttgctgctgtggaagaacctcactttcagaagaagacacagtaagcttggttttca 276

Db 345 ttgctgctgtggaagaacctcactttcagaagaagacacagtaagcttggttttca 404

QY 277 gcagcgggggttctctcattttttttgtgtgtttgagttggg 321

Db 405 gtggcctggcctctatttctcctgacccgctctgttcgg 449

## RESULT 14

AAK51683  
 ID AAK51683 standard; cDNA; 7281 BP.

XX AAK51683;

XX 06-NOV-2001 (first entry)

XX Human polynucleotide SEQ ID NO 228.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;

XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

XX tissue growth factor; immunomodulatory; cancer; leukaemia;

XX nervous system disorder; arthritis; inflammation; ss.

XX Homo sapiens.

XX WO200157190-A2.

XX 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US04098.

XX 03-FEB-2000; 2000US-0496914.

XX 27-APR-2000; 2000US-0560875.

XX 20-JUN-2000; 2000US-0598075.

XX 19-JUL-2000; 2000US-0620325.

PR 01-SEP-2000; 2000US-0654936.  
 PR 15-SEP-2000; 2000US-0663561.  
 PR 20-OCT-2000; 2000US-0693325.  
 PR 30-NOV-2000; 2000US-0728422.  
 XX (HYSE-) HYSEQ INC.  
 PA  
 XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;  
 PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;  
 XX  
 DR WPI; 2001-476283/51.  
 DR P-PSDB; AAM78550.  
 XX  
 PT Nucleic acids encoding polypeptides with cytokine-like activities,  
 XX useful in diagnosis and gene therapy -  
 PS Claim 1; Page 1086-1096; 6221pp; English.  
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the  
 CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activin/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation.  
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666  
 CC (AAM80020) are omitted as the relevant pages from the sequence listing  
 CC were missing at the time of publication.  
 XX  
 SQ Sequence 7281 BP; 1831 A; 1773 C; 1915 G; 1762 T; 0 other;

Query Match 40.8%; Score 145.8; DB 22; Length 7281;  
 Best Local Similarity 80.9%; Pred. No. 1.3e-34;  
 Matches 182; Conservative 0; Mismatches 42; Indels 1; Gaps 1;  
 QY 97 cagttatgaccagccagggcgctcctcgtcagctcgtcgcctcctccagggct 156  
 DB 247 cagttatgaccagccac-ggcgtccctcgtcgtcgtcgtcgcctcctccagggct 305  
 QY 157 cccgagccacacgctggcgctgctgctgaggaacatggcattgttgctcagctgagg 216  
 DB 306 cccgagccacacgctggcgctgctgctgaggaacatggcattgttgctcagctgagg 365  
 QY 217 ttgctgctgtggaagacacctcactttcagaagaagacacagtaagcttggttttca 276  
 DB 366 ttgctgctgtggaagacacctcactttcagaagaagacacagtaagcttggttttca 425  
 QY 277 gcagcgggggtctctcatt 321  
 DB 426 gtggcctggcctctatttctctcctcctcctcctcctcctcctcctcctcctcctcct 470

RESULT 15  
 AAS06120  
 ID AAS06120 standard; cDNA; 9741 BP.  
 XX  
 AC AAS06120;  
 XX  
 DT 12-SEP-2001 (first entry)  
 XX  
 DE Human ABC1 DNA sequence #1.  
 XX  
 KW Human; ABC1 gene; atherosclerosis; reverse transport; cholesterol;  
 KW cardiovascular; neurological; Tangier disease; LCAT deficiency;  
 KW lecithin-cholesterol acetyltransferase; malaria; diabetes; ss.  
 XX  
 OS Homo sapiens.

XX  
 FH Key Location/Qualifiers  
 FT CDS 185..6967  
 FT /\*tag= a  
 FT /product= "Human ABC1 protein"  
 XX  
 PN WO200130848-A2.  
 XX  
 XX 03-MAY-2001.  
 XX  
 XX 26-OCT-2000; 2000WO-EP10886.  
 XX  
 XX 26-OCT-1999; 99EP-0402668.  
 XX  
 XX 01-MAR-2000; 2000US-0186260.  
 XX  
 XX (AVET ) AVENTIS PHARMA SA.  
 PA  
 XX Denefle P, Rosier-Montus M, Arnould-Requigne I, Prades C, Naudin L;  
 PI Lemoline C, Duverger N, Jaye M, Searfoss GH, Remaley A, Brewer HB;  
 PI Dean M;  
 XX  
 XX WPI; 2001-316327/33.  
 DR P-PSDB; AAU02176.  
 DR  
 XX New human ABC1 nucleic acids and polypeptides for treating  
 PT atherosclerosis, malaria and diabetes -  
 XX  
 XX Claim 1; Page 204-208; 368pp; English.  
 CC The sequence represents the coding sequence #1 of human ABC1. The  
 CC nucleic acid sequence, primers and probes derived from the ABC1 sequence,  
 CC and polypeptides and vectors are useful for the prevention of  
 CC atherosclerosis, in a subject affected by a dysfunction in the reverse  
 CC transport of cholesterol. The polypeptide encoded by the ABC1 gene is  
 CC useful for screening for an active ingredient for the prevention or  
 CC treatment of a disease resulting from dysfunction in the reverse  
 CC transport of cholesterol. The nucleic acids and polypeptides are also  
 CC useful for treating and preventing cardiovascular and neurological  
 CC pathologies, and other diseases e.g. Tangier disease, lecithin-  
 CC cholesterol (LCAT) deficiency, malaria and diabetes.  
 XX  
 SQ Sequence 9741 BP; 2650 A; 2180 C; 2290 G; 2620 T; 1 other;

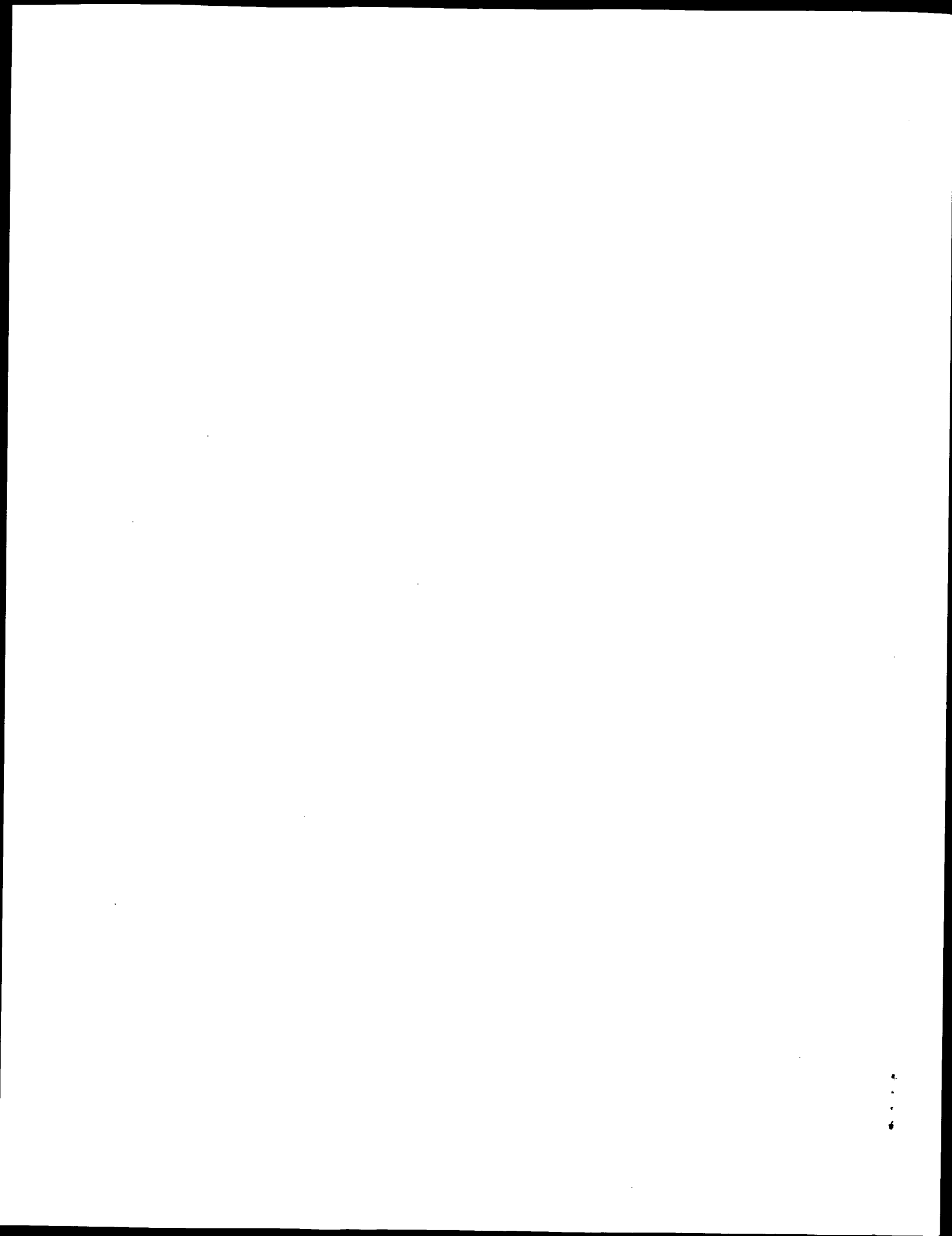
Query Match 40.8%; Score 145.8; DB 22; Length 9741;  
 Best Local Similarity 80.9%; Pred. No. 1.5e-34;  
 Matches 182; Conservative 0; Mismatches 42; Indels 1; Gaps 1;  
 QY 97 cagttatgaccagccagggcgctcctcgtcagctcgtcgcctcctccagggct 156  
 DB 90 cagttatgaccagccac-ggcgtccctcgtcgtcgtcgtcgcctcctccagggct 148  
 QY 157 cccgagccacacgctggcgctgctgctgaggaacatggcattgttgctcagctgagg 216  
 DB 149 cccgagccacacgctggcgctgctgctgaggaacatggcattgttgctcagctgagg 208  
 QY 217 ttgctgctgtggaagacacctcactttcagaagaagacacagtaagcttggttttca 276  
 DB 209 ttgctgctgtggaagacacctcactttcagaagaagacacagtaagcttggttttca 268  
 QY 277 gcagcgggggtctctcatt 321  
 DB 269 gtggcctggcctctatttctctcctcctcctcctcctcctcctcctcctcctcctcct 313

Search completed: September 20, 2002, 03:08:57  
 Job time: 10606 sec

us-09-846-456-2.std.rng

---

Fri Sep 20 08:03:38 2002





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 03:21:22 ; Search time 110.16 Seconds  
(without alignments)  
796.035 Million cell updates/sec

Title: US-09-846-456-2

Perfect score: 357

Sequence: 1 tggaggtctcagctgagagg.....gagggaaggagctgtgttg 357

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_NA:\*

- 1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*
- 2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*
- 3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*
- 4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*
- 5: /cgn2\_6/ptodata/2/ina/PTBUS\_COMB.seq:\*
- 6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	8.7	10968	2	US-08-680-327-2
2	31	8.7	10968	4	US-09-228-246-1
3	30.8	8.6	1037	1	US-08-462-126-5
4	30.8	8.6	1037	2	US-08-462-126-5
5	30.8	8.6	1037	3	US-08-636-883-3
6	30.4	8.5	2473	4	US-09-127-829-3
7	30	8.4	686	4	US-09-173-914-3
8	30	8.4	686	4	US-08-988-321B-37
9	30	8.4	4771	2	US-08-397-220B-25
10	30	8.4	4771	2	US-08-866-650-2
11	30	8.4	4771	2	US-09-021-287-2
12	30	8.4	4771	1	US-09-240-473-2
13	30	8.4	5143	1	US-08-574-043A-7
14	30	8.4	5143	2	US-08-795-015-7
15	30	8.4	7881	2	US-08-751-189-1
16	30	8.4	7881	4	US-09-060-836-1
17	30	8.4	35529	4	US-09-184-445-1
18	29.6	8.3	424	1	US-09-144-085-3
19	29.4	8.2	2480	4	US-08-609-657-15
20	29.4	8.2	3147	4	US-09-534-638-3
21	29.4	8.2	3147	4	US-08-781-802-7
22	29.4	8.2	3147	4	US-08-694-078-7
23	29.4	8.2	3147	4	US-09-058-260-7
24	29.2	8.2	31491	4	US-09-360-186-1
25	29.2	8.2	573	2	US-08-290-665A-124
26	29.2	8.2	573	5	PTBUS-095-10398-124
27	29.2	8.2	2892	1	US-08-264-534-5
28	29.2	8.2	2892	1	US-08-083-590A-1

c 28	29.2	8.2	2892	1	US-08-465-500-5	Sequence 5, Appli
c 29	29.2	8.2	2892	2	US-08-346-126-5	Sequence 5, Appli
c 30	29.2	8.2	2892	2	US-08-346-126-5	Sequence 5, Appli
c 31	29.2	8.2	2892	3	US-08-532-384-1	Sequence 1, Appli
c 32	29.2	8.2	2892	3	US-08-893-828-5	Sequence 5, Appli
c 33	28.8	8.1	1257	3	US-08-640-906-1	Sequence 1, Appli
c 34	28.8	8.1	1257	4	US-09-395-936-1	Sequence 1, Appli
c 35	28.6	8.0	2861	4	US-08-482-073-10	Sequence 10, Appli
c 36	28.6	8.0	3647	1	US-07-914-281-7	Sequence 7, Appli
c 37	28.6	8.0	3647	1	US-08-393-246-7	Sequence 7, Appli
c 38	28.6	8.0	3647	1	US-08-525-058A-7	Sequence 7, Appli
c 39	28.6	8.0	3647	2	US-08-696-731-7	Sequence 7, Appli
c 40	28.6	8.0	3647	4	US-09-042-531-7	Sequence 7, Appli
c 41	28.4	8.0	345	1	US-08-324-977-7	Sequence 7, Appli
c 42	28.4	8.0	345	2	US-08-384-616-7	Sequence 7, Appli
c 43	28.4	8.0	345	2	US-08-904-886A-7	Sequence 7, Appli
c 44	28.4	8.0	345	4	US-09-315-850-7	Sequence 7, Appli
c 45	28.4	8.0	573	2	US-08-290-665A-115	Sequence 115, App

#### ALIGNMENTS

RESULT 1  
US-08-680-327-2  
; Sequence 2, Application US/08680327  
; Patent No. 5859321  
; GENERAL INFORMATION:  
; APPLICANT: Staskawicz, Brian S., Oldroyd, Giles Edward,  
; APPLICANT: Salmeron, John M., Rommens, Caius  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PLANT  
; TITLE OF INVENTION: PATHOGEN RESISTANCE  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klarquist Sparkman Campbell Leigh &  
; ADDRESSEE: Whinston  
; STREET: One World Trade Center  
; STREET: 121 S.W. Salmon Street  
; STREET: Suite 1600  
; CITY: Portland  
; STATE: Oregon  
; COUNTRY: United States of America  
; ZIP: 97204  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Disk, 3-1/2 inch  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: MS DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/680,327  
; FILING DATE: July 11, 1996  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/310,912  
; FILING DATE: September 22, 1994  
; CLASSIFICATION: 800  
; APPLICATION NUMBER: 08/227,360  
; FILING DATE: April 13, 1994  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dow, Alan. E.  
; REGISTRATION NUMBER: 35,123  
; REFERENCE/DOCKET NUMBER: 5151-45038  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (503) 226-7391  
; TELEFAX: (503) 228-9446  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10968 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double stranded  
; TOPOLOGY: linear  
US-08-680-327-2



; Sequence 3, Application US/08636883  
; Patent No. 5830691  
; GENERAL INFORMATION:  
; APPLICANT: MIYAMURA, TATSUO  
; APPLICANT: SAITO, IZUMU  
; APPLICANT: MATSUURA, YOSHIHARU  
; APPLICANT: HONDA, YOSHIKAZU  
; APPLICANT: SEKI, MAKOTO  
; TITLE OF INVENTION: METHOD FOR PRODUCING ECTOPROTEIN OF  
; TITLE OF INVENTION: HEPATITIS C VIRUS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 S. Jefferson Davis Highway, Suite 400  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/636,883  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/446,303  
; FILING DATE:  
; APPLICATION NUMBER: US 08/074,584  
; FILING DATE: 11-JUN-1993  
; APPLICATION NUMBER: JP 152487/1992  
; FILING DATE: 11-JUN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Oblon, No. 5830691man F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 4169-003-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 413-3000  
; TELEFAX: (703) 413-2220  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1037 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Hepatitis C virus  
; IMMEDIATE SOURCE:  
; CLONE: pUCM010  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 17..1036  
US-08-636-883-3

Query Match 8.6%; Score 30.8; DB 2; Length 1037;  
Best Local Similarity 46.3%; Pred. No. 2.1;  
Matches 101; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 107 ccagccacggcgctccctgctgtagctgtggccgctgctccaggggctccagagccac 166  
DB 237 CTACCCCTCGGTTGGCGAGCCTTGGGGATAGTTGTCGCCCTCCACGAGGTTGGCGCGC 178  
QY 167 acgctgggctgtgctgtaggggaacatggcatgttgccctcagctgaggttctgtgt 226  
DB 177 TCGGAAGTCTTCCCTAGTTCGCGCGCACACCCCAACCTGGGGGCCCTTCGCGCGCAACAGGTAA 118

QY 227 ggaagaacctcacttcagaaagaagacacagtaagcttgggtttttcagcagcgggg 286  
DB 117 ACTCCACCAACGATCTGACACCGCCCGGGAACCTTAACTCTGTCGGCGCGGTTGGTG 58  
QY 287 gttctctcattttttttttgtgtgttttttgagttggggat 324  
DB 57 TTACGTTTGATTTTCTCTGGGGGTTTGGGATTGTGCT 20  
RESULT 5  
US-09-127-829-3/c  
; Sequence 3, Application US/09127829  
; Patent No. 6063904  
; GENERAL INFORMATION:  
; APPLICANT: MIYAMURA, TATSUO  
; APPLICANT: SAITO, IZUMU  
; APPLICANT: MATSUURA, YOSHIHARU  
; APPLICANT: HONDA, YOSHIKAZU  
; APPLICANT: SEKI, MAKOTO  
; TITLE OF INVENTION: METHOD FOR PRODUCING ECTOPROTEIN OF  
; TITLE OF INVENTION: HEPATITIS C VIRUS  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 S. Jefferson Davis Highway, Suite 400  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/127,829  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/446,303  
; FILING DATE:  
; APPLICATION NUMBER: US 08/074,584  
; FILING DATE: 11-JUN-1993  
; APPLICATION NUMBER: JP 152487/1992  
; FILING DATE: 11-JUN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Oblon, No. 6063904man F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 4169-003-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 413-3000  
; TELEFAX: (703) 413-2220  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1037 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Hepatitis C virus  
; IMMEDIATE SOURCE:  
; CLONE: pUCM010  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 17..1036  
US-09-127-829-3

Query Match 8.6%; Score 30.8; DB 3; Length 1037;

Best Local Similarity 46.3%; Pred. No. 2.1;		Matches 101; Conservative 0; Mismatches 117; Indels 0; Gaps 0;	
QY	107	ccagccacggcgctccctgctgtcagctctggcgcgtgctctccaggggtcccgagccac	166
Db	237	CTACCCTCGGTTGGGAGCCTTGGGATAGTGTGCGCTTCCACAGAGTTGCGCGCG	178
QY	167	acqctgggcgtgctgaggaacatggcatgttgccctcagctgaggttgcgtgt	226
Db	177	TCGGAAGTCTTCTAGTCGCGCGCACACCAACCTGGGGCCCTGGCGGCAACAGGTAA	118
QY	227	ggaagaacctcaccttcagaaagaacacagctaagcttggttttcagcagcgagg	286
Db	117	ACTCCACCAACGATCTGACACCGCCGGGAACCTTACGTCCTGTGGCGCGGTTGGTG	58
QY	287	gtctctcatTTTTCTTTTgtggttttggattggggat	324
Db	57	TTACGTTTGATTTTCTTTGGGCTTTTGGATTGTGCT	20
RESULT 6			
US-09-173-914-3/c			
; Sequence 3, Application US/09173914			
; Patent No. 6171857			
; GENERAL INFORMATION:			
; APPLICANT: Hendrickson, Eric			
; TITLE OF INVENTION: A No. 6171857el Leucine Zipper, KARP-1 and			
; FILE REFERENCE: B0877/7017/HK			
; CURRENT APPLICATION NUMBER: US/09/173,914			
; CURRENT FILING DATE: 1998-10-16			
; EARLIER APPLICATION NUMBER: 60/064,557			
; EARLIER FILING DATE: 1997-10-17			
; NUMBER OF SEQ ID NOS: 35			
; SOFTWARE: FastSeq for Windows Version 3.0			
; SEQ ID NO 3			
; LENGTH: 2473			
; TYPE: DNA			
; ORGANISM: Mesocricetus Auratus			
US-09-173-914-3			
Query Match 8.5%; Score 30.4; DB 4; Length 2473;		Best Local Similarity 57.3%; Pred. No. 4.5;	
Matches 55; Conservative 0; Mismatches 41; Indels 0; Gaps 0;			
QY	250	agacaacagtaagctgtgtttttcagcagcggggttctctcattttcttggg	309
Db	1629	AGAGGAATCTCATTTGGCTTTTCACCTCAGTGGGGGATCCAGCATATTTCAAAATGTGC	1570
QY	310	ttttgattgggattggaggaggaggaggaggag	345
Db	1569	TGCTGAATTGGGGGTAGACGCTCTCTGGGATGTAAG	1534
RESULT 7			
US-08-988-321B-37/c			
; Sequence 37, Application US/08988321B			
; Patent No. 6174868			
; GENERAL INFORMATION:			
; APPLICANT: Kevin P. Anderson et al.			
; TITLE OF INVENTION: Compositions And Methods For Treatment Of Hepatitis C v			
; NUMBER OF SEQUENCES: 37			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Law Offices of Jane Massey Licata			
; STREET: 66 East Main Street			
; CITY: Marlton			
; STATE: NJ			
; COUNTRY: USA			
; ZIP: 08053			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE			
; COMPUTER: IBM COMPATIBLE			

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/866,650
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Berson, Bennett J
; REGISTRATION NUMBER: 37094
; REFERENCE/DOCKET NUMBER: 960296.93839
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4771 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 611..3652
; OTHER INFORMATION: /product= "murine mTll prot
; PS-08-866-650-2

```

TOPOLOGY: Linear  
 ANTI-SENSE: No  
 SEQUENCE DESCRIPTION: SEQ ID NO: 25:  
 S-08-397-220B-25  
 Query Match 8.4%; Score 30; DB 4; Length 686;  
 Best Local Similarity 46.3%; Pred. No. 3.1;  
 Matches 99; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

Qy	102	aatgacacagccaggcgctccctctgctcagctctgacgctgccttcacagggctccga	161
Db	141	ATTATATCTTCCCGACACACGGCGGTGGACAGCGGGCGGCAGCGGCTGCAGCTCGCTC	200
Qy	162	gccacacgctggcgctgctgctgtagggagaacatgycatgttgctcgaactagaggttgt	221
Db	201	TCGGCCCGGGGTCCTACACAGCGCGGGCGCGCGCGGAGCCGGAGCTCCGGTGGCA	260
Qy	222	gctgtg	227
Db	261	GCTGAG	266

```

RESULT 10
US-09-021-287-2
; Sequence 2, Application US/09021287
; Patent No. 5981717
; GENERAL INFORMATION:
; APPLICANT: Greenspan, Daniel S
; APPLICANT: Takahara, Kazuhiko
; APPLICANT: Hoffman, Guy G
; TITLE OF INVENTION: Mammalian Tolloid-Like Protein
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/021,287
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/866,650

```

```
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Berson, Bennett J
; REGISTRATION NUMBER: 37094
; REFERENCE/DOCKET NUMBER: 960296.93839
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4771 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 611..3652
; OTHER INFORMATION: /product= "murine mtll protein"
; US-09-021-287-2

Query Match      8.4%; Score 30; DB 2; Length 4771;
Best Local Similarity 52.4%; Pred. No. 8.5;
Matches 66; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 102 aatgaccagcagcggtccctgtctgtcagctctgtgcccgtctgtccctccagggctcccca 161
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 141 ATTCACTTCCCGGAGCAGCGCGGTGGCAGCGCGGCGCGGCGCGGCGGTGCAGCTCGGCTC 200

QY 162 gccacacgctggcgctgctgagtgaggaacatggcatgttgcctcagctgaggttgc 221
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 201 TCGGCGCGGGGTCTTACAGCGCGCGGGGGCGGCGCGGCGGAGCGGAGCTCCGGTGCCA 260

QY 222 gctgtg 227
   |||||
Db 261 GCTGAG 266

RESULT 11
US-09-240-473-2
; Sequence 2, Application US/09240473
; Patent No. 6297011
; GENERAL INFORMATION:
; APPLICANT: Greenspan, Daniel S
; APPLICANT: Takahara, Kazuhiko
; APPLICANT: Hofman, Guy G
; TITLE OF INVENTION: Mammalian Tolloid-Like Protein
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/240.473
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Berson, Bennett J
; REGISTRATION NUMBER: 37094
; REFERENCE/DOCKET NUMBER: 960296.93839
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4771 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4771 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 611..3652
; OTHER INFORMATION: /product= "murine mtll protein"
; US-09-240-473-2

Query Match      8.4%; Score 30; DB 4; Length 4771;
Best Local Similarity 52.4%; Pred. No. 8.5;
Matches 66; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 102 aatgaccagcagcggtccctgtctgtcagctctgtgcccgtctgtccctccagggctcccca 161
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 141 ATTCACTTCCCGGAGCAGCGCGGTGGCAGCGCGGCGCGGCGGTGCAGCTCGGCTC 200

QY 162 gccacacgctggcgctgctgagtgaggaacatggcatgttgcctcagctgaggttgc 221
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 201 TCGGCGCGGGGTCTTACAGCGCGCGGGGGCGGCGCGGAGCGGAGCTCCGGTGCCA 260

QY 222 gctgtg 227
   |||||
Db 261 GCTGAG 266

RESULT 12
US-08-574-043A-7/c
; Sequence 7, Application US/08574043A
; Patent No. 5807692
; GENERAL INFORMATION:
; APPLICANT: Kinzler, Kenneth W.
; APPLICANT: El-Deiry, Wafik
; APPLICANT: Vogelstein, Bert
; TITLE OF INVENTION: p21WAF1 Derivatives and Diagnostic
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Allegretti, LTD
; STREET: 1001 G Street, NW suite 1100
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/574,043A
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 01107.49698
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202.508.9100
; TELEFAX: 202.508.9299
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5143 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
```



Db 5166 TGCTGTTTGTTCCTCTCCGATGATACACTCTTTCTTACTGCTTCGACGGGCTCCTGGA 5225  
QY 126 cgtgcagctctgcccgtgcctccagggctcccgagccacacgct 171  
Db 5226 GCTCTGGGACCTGCAGCATGGTTGTCGGGTGCTGCAGACTAAGGCT 5271

## RESULT 15

US-09-060-836-1  
; Sequence 1, Application US/09060836  
; Patent No. 5981707  
; GENERAL INFORMATION:  
; APPLICANT: Harrington, Lea A.  
; APPLICANT: Robinson, Murray O.  
; TITLE OF INVENTION: No. 5981707el Genes Encoding Telomerase Protein  
; TITLE OF INVENTION: 1  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Amgen, Inc.  
; STREET: 1840 De Havilland Drive  
; CITY: Thousand Oaks  
; STATE: California  
; COUNTRY: USA  
; ZIP: 91320-1789  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/060,836  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/751,189  
; FILING DATE: 15-NOV-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Oleski, Nancy A.  
; REGISTRATION NUMBER: 34,688  
; REFERENCE/DOCKET NUMBER: A-433  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7881 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; US-09-060-836-1

Query Match 8.4%; Score 30; DB 2; Length 7881;  
Best Local Similarity 48.8%; Pred. No. 11;  
Matches 81; Conservative 0; Mismatches 85; Indels 0; Gaps 0;  
QY 6 gctcagctgagagggctgattagcagctcctcattggtgtagcttgcagcaataac 65  
Db 5106 GTTGGACCTCAGAACTGGCAGGAGGAGAGTCTGTGGTGTGAGTGGCTGTGATGGAATCTC 5165  
QY 66 tgatggctgtttccctcctgctttatcttttaagttaatgaccagccagggcggtccctg 125  
Db 5166 TGCTGTTTGTTCCTCTCCGATGATACACTCTTTCTTACTGCTTCGACGGGCTCCTGGA 5225  
QY 126 ctgtcagctctgcccgtgcctccagggctcccgagccacacgct 171  
Db 5226 GCTCTGGGACCTGCAGCATGGTTGTCGGGTGCTGCAGACTAAGGCT 5271

Search completed: September 20, 2002, 03:21:31  
Job time: 14180 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 01:23:10 ; Search time 3895 Seconds  
(without alignments)  
1237.076 Million cell updates/sec

Title: US-09-846-456-2  
Perfect score: 357  
Sequence: 1 tggaggtctcagctgagagg.....gagggaaggaagctgtgttg 357

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_estl:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	144.2	40.4	736	9	AU135588
2	75.8	21.2	619	9	BB657864
3	73.8	20.7	292	10	Z44377 HSC1ZB081 n
4	71.2	19.9	535	10	BG384217 303216 MA
5	42.4	11.9	982	12	AL192682 Tetraodon
6	39.6	11.1	910	12	CNS0060N
7	37.4	10.5	926	12	AZ542175 ENTGP86TF
8	37	10.4	440	9	AA914462 v201f08.r
9	37	10.4	533	9	AV845237 AV845237
10	36.4	10.2	571	9	AV864011 AV864011
11	36.4	10.2	579	9	AV892280 AV892280
12	36	10.1	322	9	AU056364 AU056364
13	36	10.1	393	9	AU162694 AU162694
14	36	10.1	571	9	AV862022 AV862022
15	36	10.1	872	12	CNS027E7
16	35.8	10.0	344	10	BM149133 TCAAP2E63
17	35.8	10.0	785	12	AF010859 AF010859

C 18	35.8	10.0	939	12	CNS00CNG
19	35.6	10.0	514	12	BH087817
C 20	35.6	10.0	1097	12	CNS04PHC
21	35.6	10.0	1574	10	BG325753
22	35.4	9.9	390	9	BB841907
23	35.4	9.9	482	12	AQ848626
C 24	35.4	9.9	937	12	CNS006ST
C 25	35.4	9.9	1901	10	BF128237
C 26	35.2	9.9	972	10	BE733920
27	35.2	9.9	1201	12	CNS0162E
C 28	34.8	9.7	301	9	AW751638
29	34.8	9.7	365	9	BB843149
30	34.8	9.7	516	10	BF484412
31	34.8	9.7	530	10	BE471178
C 32	34.8	9.7	810	10	BE748387
33	34.8	9.7	1068	12	CNS00ETV
34	34.6	9.7	483	12	CNS03210
35	34.6	9.7	884	12	CNS006U0
36	34.6	9.7	1039	12	CNS015P6
C 37	34.6	9.7	1068	12	CNS00ETV
38	34.6	9.7	1101	12	CNS00LT2
39	34.4	9.6	460	9	AU212427
C 40	34.4	9.6	778	10	BI886434
C 41	34.2	9.6	619	9	AW681359
42	34.2	9.6	896	12	CNS00BP8
43	34.2	9.6	970	12	CNS010C9
C 44	34.2	9.6	1201	12	CNS01604
C 45	34	9.5	468	12	AQ773862

## ALIGNMENTS

RESULT 1

AU135588  
LOCUS AU135588 PLAC1 Homo sapiens cDNA clone PLACE1002437 5', mRNA linear EST 24-Oct-2000  
DEFINITION AU135588.1 GI:10996127  
ACCESSION AU135588  
VERSION AU135588.1  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1 (bases 1 to 736)  
AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and Isogai,T.  
TITLE HRI human cDNA project  
JOURNAL Unpublished (2000)  
COMMENT Contact: Takao Isogai  
Genomics Laboratory  
Helix Research Institute  
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
Tel: 81-438-52-3951  
Fax: 81-438-52-3952  
Email: genomics@hri.co.jp  
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.  
FEATURES  
Location/Qualifiers  
1..736  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="PLACE1002437"  
/clone\_lib="PLACE1"  
/tissue\_type="placenta"  
/notes="Vector: pME18SFL3"

BASE COUNT 163 a 199 c 199 g 170 t 5 others  
ORIGIN

Query Match 40.4%; Score 144.2; DB 9; Length 736;  
 Best Local Similarity 80.4%; Pred. No. 8.3e-28;  
 Matches 181; Conservative 0; Mismatches 43; Indels 1; Gaps 1;

QY 97 cagtaataaccagcagcggcctcctctgctcaagctggtccgctgctccagggt 156  
 |||||  
 Db 219 CAGTTAATGACCAACCCAC-GGGCTCCCTGCTGTGAGCTGTGGCCGCTTCCAGGGCT 277  
 |||||  
 QY 157 ccgagcacacacactggcgtgctgaggaacatgcatgttggcctcagctgagg 216  
 |||||  
 Db 278 CCCAGCACACGCTGGGGTGTGCTGAGGAAACATGCTGTGGCCCGCAGCTGAGG 337  
 |||||  
 QY 217 ttctgctgtggaagacactcaatttcagaagaagcaaacagtaagcttgggttttca 276  
 |||||  
 Db 338 TTCTGCTGTGTGGAAGAACCTCACTTTTCAAGAAGAACAAACATGTGCTGCTGGAA 397  
 |||||  
 QY 277 gcagcgggggtctctcatttttttttttgggttttgggtggg 321  
 |||||  
 Db 398 GTGGCTGGCTCTATTATCTTCGTGACCTGATCTGTGTCGG 442

## RESULT 2

BB657864 BB657864 619 bp mRNA linear EST 26-OCT-2001  
 LOCUS BB657864 RIKEN full-length enriched, 12 days embryo eyeball Mus  
 DEFINITION musculus cDNA clone D230019D04 5', mRNA sequence.  
 ACCESSION BB657864  
 VERSION BB657864.1 GI:16491690  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 AUTHORS 1 (bases 1 to 619)

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,  
 Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda  
 , M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M.,  
 Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki  
 , D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,  
 Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T.,  
 Muramatsu, M. and Hayashizaki, Y.  
 RIKEN Mouse ESTs (Arakawa, T., et al. 2001)  
 Unpublished (2001)

## TITLE

## JOURNAL

## COMMENT

Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216

Email: genome-res@gsc.riken.go.jp,  
 URL: http://genome.gsc.riken.go.jp/  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh  
 , M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
 Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura  
 , S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and  
 Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10  
 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara  
 , Y. and Hayashizaki, Y.  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa  
 , K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and  
 Hayashizaki, Y.

Computational Analysis of Full-Length Mouse cDNAs Compared with  
 Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)  
 Please visit our web site (http://genome.gsc.riken.go.jp) for  
 further details.

## FEATURES

## source

Location/Qualifiers  
 1..619  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="D230019D04"  
 /clone\_lib="RIKEN full-length enriched, 12 days embryo  
 eyeball"  
 /tissue\_type="eyeball"  
 /dev\_stage="12 days embryo"  
 /lab\_host="DH10B"  
 /note="Site 1: SalI; Site 2: BamHI; cDNA library was  
 prepared and sequenced in Mouse Genome Encyclopedia  
 Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in  
 RIKEN. Division of Experimental Animal Research in Riken  
 contributed to prepare mouse tissues. 1st strand cDNA was  
 primed with a primer [5'  
 GAGAGAGAGCGCGCGCAACTCGAGTTTTTTTTTTTTTTTNN 3'], cDNA was  
 prepared by using trehalose thermo-activated reverse  
 transcriptase and subsequently enriched for full-length by  
 cap-trapper. Second strand cDNA was prepared with the  
 primer adapter of sequence [5'  
 GAGAGAGAGATTCTCGATTAAATTAATCCGCCCCCCC 3']. cDNA  
 was cleaved with BamHI and XhoI. Vector: a modified  
 pBluescript KS(+) after bulk excision from Lambda FIC I."  
 BASE COUNT 126 a 185 c 173 g 134 t 1 others  
 ORIGIN

Query Match 21.2%; Score 75.8; DB 9; Length 619;  
 Best Local Similarity 77.8%; Pred. No. 1.1e-09;  
 Matches 133; Conservative 0; Mismatches 27; Indels 11; Gaps 3;  
 QY 97 cagtaataaccagcagcggcctcctctgctcaagctggtccgctgctccagggt 156  
 |||||  
 Db 219 CAGTTAATGACCAACCCAC-AGAGTCACAGCTCTGTGCTGTGGCTGCT 276  
 |||||  
 QY 157 cccagagccaca-----cgctggcgctgctgctgaggaacatggttggcct 207  
 |||||  
 Db 277 CTCGAGCGCGCAGCAGCGAGTCGCTGTGGTGGCGCTGTGTGACATGGTGTGGCCT 336  
 |||||  
 QY 208 cagctgaggttgcctgtggaagaaacctcacttccagaagaacacaaaca 258  
 |||||  
 Db 337 CAGTTAAGGCTCTGCTGTGGAGAAATCTGACATTCGAGGAGACAAACA 387

## RESULT 3

244377 3  
 LOCUS 244377 292 bp mRNA linear EST 14-NOV-1994  
 DEFINITION HSC12B081 normalized infant brain cDNA Homo sapiens cDNA clone  
 C-12B081, mRNA sequence.  
 ACCESSION 244377  
 VERSION 244377.1 GI:573506  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 292)

Auffray, C., Behar, G., Bois, F., Bouchier, C., da Silva, C., Devignes  
 , M.D., Duprat, S., Houlgatte, R., Jumeau, M.N., Lamy, B., Lorenzo, F.,  
 Mitchell, H., Mariage-Samson, K., Pietu, G., Pouliot, Y.,  
 Sebastiani-Kabatchis, C. and Tessier, A.  
 IMAGE: molecular integration of the analysis of the human genome  
 and its expression

JOURNAL C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)  
 MEDLINE 95277534  
 COMMENT Contact: Genethon



```

/db_xref="taxon:99883"
/clone="260p19"
/clone_lib="G"
/note="Genoscope sequence ID : COAG260CH10SP1-end :
PUC-ori"
BASE COUNT      246 a  248 c  224 g  247 t  17 others
ORIGIN

Query Match      11.9%; Score 42.4; DB 12; Length 982;
Best Local Similarity 55.0%; Pred. No. 0.86;
Matches 66; Conservative 8; Mismatches 46; Indels 0; Gaps 0;

QY 237 cactttcagaagaacacgtagcttggttttttcagcagcggtttctctcat 296
||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
DB 839 CACACGACATRAAKAWAYTTTCAGACCGGCTTBTCTGSRGKGDDGGTCTTTT 780
||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

QY 297 tttttcttggttttgattggattggaggaggaggaggaggaggaggagctgttt 356
||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
DB 779 TTTTNTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

RESULT 6
CNS0060N      910 bp  DNA  linear  GSS 03-JUN-1999
LOCUS      Drosophila melanogaster genome survey sequence T7 end of BAC #
DEFINITION  BACR14J21 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION  AL065629
VERSION    AL065629.1 GI:4944698
KEYWORDS   GSS:
SOURCE     fruit fly.
ORGANISM   Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
            1 (bases 1 to 910)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osogawa and
Aaron Mammoss in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             source
Location/Qualifiers
1..910
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RPCI-98"
/clone="BACR14J21"
/note="end : T7"
BASE COUNT      202 a  63 c  112 g  198 t  335 others
ORIGIN

Query Match      11.18; Score 39.6; DB 12; Length 910;
Best Local Similarity 50.8%; Pred. No. 4.7;
Matches 64; Conservative 11; Mismatches 51; Indels 0; Gaps 0;

QY 229 aagaacctcacttcagaagaacacagtaagcttggttttttcagcagcggtgt 288
||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

```

```

Db 369 AAAAAAAAAAATWAAAAAAAAAAAAAAAAAKARWGGKAGCGGTGAGCGCDTGTGTGTT 428
||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 289 tctctcattttttttgtggtttttgtggtttttgtggtttttgtggtttttgtggttt 348
||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 429 TTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 488
||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 349 gctgtg 354
||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 489 GGGGG 494
||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

```

```

RESULT 7
LOCUS      AZ542175      926 bp  DNA  linear  GSS 14-NOV-2000
DEFINITION  ENTGP86TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
genomic, DNA sequence.
ACCESSION  AZ542175
VERSION    AZ542175.1 GI:11150635
KEYWORDS   GSS:
SOURCE     Entamoeba histolytica.
ORGANISM   Entamoeba histolytica
            Eukaryota; Entamoebidae; Entamoeba.
            1 (bases 1 to 926)
REFERENCE  1 (bases 1 to 926)
AUTHORS   Loftus,B., Van Aken,S. and Fraser,C.
TITLE     Determination of clone end sequences from Entamoeba histolytica
          HM1:IMSS sheared DNA library
JOURNAL   Unpublished (2000)
COMMENT   Contact: Brendan J Loftus
          Department of Eukaryotic Genomics
          The Institute for Genomic Research
          9712 Medical Center Dr., Rockville, MD 20850, USA
          Tel: 301 838 0208
          Fax: 301 838 3543
          Email: bjloftus@tigr.org
          Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
          DNA library
          Seq primer: M13-Forward
          Class: shotgun
          High quality sequence start: 17
          High quality sequence stop: 297.
FEATURES             Location/Qualifiers
1..926
/organism="Entamoeba histolytica"
/strain="HM1:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/note="Vector: PHOS1; Site1: Bst I; Constructed at The
Institute for Genomic Research (IGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barell, Oxford University Press, 1999)."
BASE COUNT      304 a  29 c  219 g  374 t
ORIGIN

```

```

Query Match      10.5%; Score 37.4; DB 12; Length 926;
Best Local Similarity 57.1%; Pred. No. 18;
Matches 68; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 228 gaagaaccttcacttcagaagaacacagtaagcttggttttttcagcagcggtgt 287
||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 757 GAAAAAGGGAAGAAAAAGAAAAAGAAAAAGAAAGTTGTTTGTGGGGGGTGGG 816
||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 288 tctctcattttttttgtggtttgtggtttgtggtttgtggtttgtggtttgtggttt 346
||||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

```



\_\_\_\_\_

```

VERSION      AU162694.1  GI:11026093
KEYWORDS
SOURCE       Oryza sativa
ORGANISM     Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
              Ehrhartoideae; Oryzeae; Oryza.
REFERENCE    1 (bases 1 to 393)
AUTHORS      Sasaki, T. and Yamamoto, K.
TITLE        Rice cDNA from mature leaf (2000)
JOURNAL      Unpublished (2000)
COMMENT      Contact: Takuji Sasaki
              National Institute of Agrobiological Resources
              Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
              305-8602, Japan
              Tel: 81-298-38-7441
              Fax: 81-298-38-7468
              Email: tsasaki@abr.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/
              PROJECT = 'RGP'.
              S20642_1A.

FEATURES
  source
    1..393
    /organism="Oryza sativa"
    /strain="Nipponbare"
    /db_xref="taxon:4530"
    /clone="S20642"
    /clone_lib="Rice mature leaf"
    /tissue_type="mature leaf"
    /dev_stage="egg"

BASE COUNT  54 a 107 c 146 g 86 t
ORIGIN

Query Match      10.1%; Score 36; DB 9; Length 393;
Best Local Similarity 53.6%; Pred. No. 38;
Matches 75; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY  213 gaggtgtgctgtggaagaacctcaatttcagaagaagacaacagtaagcttggttt 272
DB  95 GGGGAGGAGCGGTGGTGGTCCCGTCGGGAGTTTCGATTGGTTGGTGGTGA 154
QY  273 ttacgacgggggggtctctcattttcttctgtggtttgagttgggagttggagag 332
DB  155 GTTGTGCTGGGTGGTGGGAGTTGGTTGGTGGAGTTGGAGATGGATTTCCGGGAG 214
QY  333 ggaggaggggaagagctg 352
DB  215 GGAGCGGAGGCGAGCAGCAG 234

RESULT 14
LOCUS      AV862022/c
DEFINITION AV862022 Nori Satoh unpublished cDNA library, egg Ciona
            intestinalis cDNA clone rcieg3lo05 3', mRNA sequence.
ACCESSION  AV862022
VERSION     AV862022.1  GI:16849546
KEYWORDS
SOURCE      Ciona intestinalis.
ORGANISM    Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
              Phlebobranchia; Cionidae; Ciona.
REFERENCE    1 (bases 1 to 571)
AUTHORS      Satoh, N., Satou, Y., Kohara, Y. and Shin-i, T.
TITLE        Expressed genes in Ciona intestinalis
JOURNAL      Unpublished (2000)
COMMENT      Contact: Nori Satoh
              Department of Zoology
              Kyoto University
              Sakyo-ku, Kyoto 606-8502, Japan
              Tel: 81-75-753-4081
              Fax: 81-75-705-1113
              Email: satoh@ascidian.zool.kyoto-u.ac.jp.
              Location/Qualifiers

FEATURES
  source
    1..571
    /organism="Ciona intestinalis"
    /db_xref="taxon:7719"
    /clone="rcieg3lo05"
    /clone_lib="Nori Satoh unpublished cDNA library, egg"
    /tissue_type="whole animal"
    /dev_stage="egg"

BASE COUNT  154 a 116 c 99 g 201 t 1 others
ORIGIN

Query Match      10.1%; Score 36; DB 9; Length 571;
Best Local Similarity 50.9%; Pred. No. 40;
Matches 84; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY  98 agttaatgaccagcagggcgctccctgtgtcagctctgtgcccgtctccagggctc 157
DB  216 AGAATATGGAGAGCCCAAGATCTGCATGTTCTGTACACAGAACTCGCTTTCATTGGATC 157
QY  158 ccgagccacacgctggcgctggtggcgaggaacatggcctggtggtcagctgaggt 217
DB  156 AAAATGCCAACGCTGTGCCAGCAGTAAGAAGAAATATGAATGCCAGTCAATTGTGAGAA 97
QY  218 tgcgtgtggaagaacctcaatttcagaagaagacaacagtaa 262
DB  96 ATGTAATTTGAAAGCAGCATTTTATCATCAAAAAAGCTCATATTGA 52

RESULT 15
LOCUS      CNS027E7
DEFINITION Tetraodon nigroviridis genome survey sequence PUC-ori end of clone
            242P21 of library G from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION  AL184552
VERSION     AL184552.1  GI:7822656
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis.
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
              Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
              Tetraodontidae; Tetraodon.
REFERENCE    1 (bases 1 to 872)
AUTHORS      Roest-Crollius, H., Jaillon, O., Dasilva, C., Fizames, C., Fisher, C.,
              Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and
              Weissenbach, J.
TITLE        Characterization and repeat analysis of the compact genome of the
              freshwater pufferfish Tetraodon nigroviridis
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 872)
AUTHORS      Roest-Crollius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,
              Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F.,
              Saurin, W. and Weissenbach, J.
TITLE        Human gene number estimate provided by genome wide analysis using
              Tetraodon nigroviridis DNA sequence
JOURNAL      Unpublished
REFERENCE    3 (bases 1 to 872)
AUTHORS      Genoscope.
TITLE        Direct Submission
JOURNAL      Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
COMMENT      This sequence is a single read and was generated as part of a large
              scale clone-end sequencing project of the Tetraodon nigroviridis
              genome. For more information, please take a look at
              http://www.genoscope.cns.fr/Tetraodon.
              Location/Qualifiers

FEATURES
  source
    1..872
    /organism="Tetraodon nigroviridis"
    /db_xref="taxon:99883"
    /clone="242P21"
    /clone_lib="G"
    /note="Genoscope sequence ID : COAG242CH1ISP1-end ;
           PUC-ori"

```

BASE COUNT	377 a	61 c	186 g	179 t	69 others
ORIGIN					

Query Match 10.1%; Score 36; DB 12; Length 872;  
Best Local Similarity 55.3%; Pred. No. 42;  
Matches 63; Conservative 3; Mismatches 48; Indels 0; Gaps 0;

Qy	244	agaagacaaacagtaagcttgggtttttcagcagcgggggtctctcatcttctt	303
		_ _	
Db	571	AAAAAAAAAAGAGRAARGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGTTT	630

Oy	304	ttgtgggttttgagttgggattggaggagggaggaggggaagaagctgtgttg	357
		: :	
Dd	631	TTTTTTTTTTTTTTTTTTGGGGGGGGGGGGGGGGGGGKGGGKKR	684

Search completed: September 20, 2002, 01:23:14  
Job time: 7374 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 03:13:43 ; Search time 5225.75 Seconds  
(without alignments)  
884.995 Million cell updates/sec

Title: US-09-846-456-4

Perfect score: 221

Sequence: 1 gtaattgcgcgcgagtgta.....aacacaaagtggaaaacag 221

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pi.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vi.\*

30: em\_htg\_hum.\*

31: em\_htg\_inv.\*

32: em\_htg\_other.\*

33: em\_htgo\_inv.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
-----					

1	221	100.0	221	6	AX351032	AX351032 Sequence
2	221	100.0	697	9	AF258627	Homo sapi
3	221	100.0	1167	9	AJ252201	Homo sapi
4	221	100.0	1167	9	AF258623	Homo sapi
5	221	100.0	3231	6	AF258623S1	Sequence
6	221	100.0	7260	6	AX351029	Sequence
7	221	100.0	96717	9	AX253452	Sequence
8	221	100.0	149034	9	AL359182	Human DNA
9	221	100.0	149034	2	AF275948	Homo sapi
10	221	100.0	201144	2	AC012230	Homo sapi
11	220.6	99.8	183999	6	AF287262	Homo sapi
12	220	99.5	69570	2	AX092589	Sequence
13	219.4	99.3	1556	2	AC021246	Homo sapi
14	219	99.1	1750	9	AK024328	Homo sapi
15	214.8	97.2	90698	2	AK022254	Homo sapi
16	205	92.8	9854	6	AC021345	Homo sapi
17	205	92.8	9854	6	AX127831	Sequence
18	203.4	92.0	69570	2	AX139818	Sequence
19	197	89.1	10442	6	AC021246	Homo sapi
20	197	89.1	10442	6	AX060713	Sequence
21	197	89.1	10442	9	AX060892	Sequence
22	190	86.0	10474	6	AF285167	Homo sapi
23	190	86.0	10474	6	AX060719	Sequence
24	190	86.0	10474	6	AX060898	Sequence
25	190	86.0	10474	6	AX060900	Sequence
26	92	41.6	446	6	AX127764	Sequence
27	92	41.6	446	6	AX127764	Sequence
28	92	41.6	9741	6	AX139751	Sequence
29	92	41.6	9741	6	AX127830	Sequence
30	92	41.6	9741	6	AX139817	Sequence
31	91	41.2	1643	6	AX351038	Sequence
32	91	41.2	1643	6	AX060715	Sequence
33	67.4	30.5	278572	10	AX060894	Sequence
34	39.2	17.7	54705	2	AF287263	Mus muscu
35	38.2	17.3	35122	9	AC103155	Rattus no
36	38.2	17.3	100000	9	AP000302	Homo sapi
37	38.2	17.3	100000	9	AP000046	Homo sapi
38	38.2	17.3	100000	9	AP000114	Homo sapi
39	38.2	17.3	100000	9	AP000190	Homo sapi
40	38.2	17.3	187272	2	AC096044	Homo sapi
41	37.2	16.8	13613	9	AP001717	Rattus no
42	36.6	16.6	204924	2	AC093942	Rattus no
43	35.6	16.1	2344	10	AC015616	Homo sapi
44	35.6	16.1	2674	10	MMVGR6	U73520 Mus musculu
45	35.4	16.0	110000	2	BC016252	Mus muscu
						Continuation (19 0

## ALIGNMENTS

RESULT	1	AX351032	Sequence 4 from Patent WO0183746.	221 bp	DNA	linear	PAT 06-FEB-2002
LOCUS	AX351032	Sequence 4 from Patent WO0183746.					
DEFINITION	AX351032	Sequence 4 from Patent WO0183746.					
ACCESSION	AX351032	Sequence 4 from Patent WO0183746.					
VERSION	AX351032.1	GI:18616388					
KEYWORDS	human.						
SOURCE	human.						
ORGANISM	Homo sapiens						
REFERENCE	1 (sites)						
AUTHORS	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.						
TITLE	Rosier-Montus M.F., Prades, C., Lemoine, C., Naudin, L., Deneffe, P., Brewer, B., Duverger, N., Renaley, A. and Santamarina-Fojo, S.						
JOURNAL	Regulatory nucleic acid sequences of the abcl gene						
FEATURES	Patent: WO 0183746-A 4 08-NOV-2001; Aventis Pharma S.A. (FR)						
source	Location/Qualifiers						
	1..221						
	/organism="Homo sapiens"						
	/db_xref="taxon:9606"						
BASE COUNT	44 a	62 c	73 g	42 t			
ORIGIN							

```

Query Match      100.0%; Score 221; DB 6; Length 221;
Best Local Similarity 100.0%; Pred. No. 5.4e-49;
Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gtaattgcgagcagagtgagtgaggccgagcccgagccgagccgagcccttctc 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 GTAATTGCGAGCAGAGTGAGTGGGCGCGGAGCCGCGAGAGCCGAGCCGCTTCTCTC 60

QY 61 ccgggctcgcgcagcagcagcagcagcagcagcagcagcagcagcagcagc 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 CCGGCTCGCGGCGAGGAGCGGCGGAGCTCCGCGCACCACAGAGCCGCTTCTCAGGCG 120

QY 121 gcttgctcctctgtttttcccggttctgttttccctctctcgcgaagcgtgtcaa 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 GCITTTGCTCTGTTTTCCTCCGCTTCTGTTTTCCTCCGCTTCTCAGGCGTGTCAA 180

QY 181 ggggtgaggaagagacgacacacacacacacacacacacacacacacacacag 221
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 181 GGGGTAGGAGAGAGACGACCAACACACAAAGTGGAAGACAG 221

RESULT 2
AF258627
LOCUS      AF258627 Homo sapiens ATP binding cassette transporter 1 (ABCA1) mRNA, partial cds. 697 bp mRNA linear PRI 11-MAY-2000
DEFINITION
ACCESSION AF258627
VERSION AF258627.1 GI:7769707
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 697)
AUTHORS Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C., Aouizerat,B.E., Fielding,C.J. and Kane,J.P.
TITLE Analysis of hABC1 gene 5' end: additional peptide sequence, promoter region, and four polymorphisms
JOURNAL Biochem. Biophys. Res. Commun. 271 (2000) In press
REFERENCE 2 (bases 1 to 697)
AUTHORS Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C., Aouizerat,B.E., Fielding,C.J. and Kane,J.P.
TITLE Direct Submission
JOURNAL Submitted (19-APR-2000) Cardiovascular Research Institute, University of California, San Francisco, 505 Parnassus Avenue, San Francisco, CA 94143-0130, USA
FEATURES
source
1..697
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/map="9q31"
/tissue_type="placenta"
1..>697
/gene="ABCA1"
/feature="ABC1"
396..>697
/feature="membrane-bound"
/codon_start=1
/product="ATP binding cassette transporter 1"
/protein_id="AAF69513.1"
/db_xref="GI:7769708"
/translation="WACWPLELLWKNLTPRRCTCOLLLELVAMPFLIFLILSVRL
STPTEQHCHPFKAMPSAGTLFWQGLICNANPCFRPTPGAPGVGNFNS"
BASE COUNT 152 a 198 c 190 g 156 t 1 others
ORIGIN

Query Match      100.0%; Score 221; DB 9; Length 697;
Best Local Similarity 100.0%; Pred. No. 5.7e-49;
Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gtaattgcgagcagagtgagtgaggccgagcccgagccgagccgagcccttctc 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 927 GTAATTGCGAGCAGAGTGAGTGGGCGCGGAGCCGCGAGAGCCGAGCCGCTTCTCTC 60

```

Db	987	CCGGGCTCGCGCAGGGCAGGGCGGGGAGCTCCGCGCACCAACAGACGGCGGTCTTCAGGGC	1046
QY	121	gctttgctctctgtttttcccggttctgttttccctcttccggaaggttgtaa	180
Db	1047	GCATTGCTCTCTGTTTTCCTCCGGTCTGTTTCTCCCTCTCCGGAAGGCTTGTCAA	1106
QY	181	ggggtagagaaagagacgcaaacacaaaagtggaaaacag	221
Db	1107	GGGCTAGGAGAAAGAGACGCAACACAAAAGTGGAACAG	1147
RESULT	5		
AX351029			
LOCUS	AX351029	3231 bp	DNA linear PAT 06-FEB-2002
DEFINITION	Sequence 1 from Patent WO0183746.		
ACCESSION	AX351029		
VERSION	AX351029.1	GI:18616385	
KEYWORDS	human.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (sites)		
AUTHORS	Rosier-Montus, M.F., Prades, C., Lemoine, C., Naudin, L., Deneffe, P., Brewer, B., Duverger, N., Remaley, A. and Santamarina-Fojo, S.		
TITLE	Regulatory nucleic acid sequences of the abcl gene		
JOURNAL	Patent: WO 0183746-A 1 08-NOV-2001;		
FEATURES	Avantis Pharma S.A. (FR)		
source	Location/Qualifiers		
	1..3231		
BASE COUNT	809 a 773 c 876 g 773 t		
ORIGIN	/organism="Homo sapiens" /db_xref="taxon:9606"		
Query Match	100.0%; Score 221; DB 6; Length 3231;		
Best Local Similarity	100.0%; Pred. No. 6.3e-49;		
Matches	221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	gtaattcgacgagagtgagtgggcgagccgcagagccgagccgacctctc	60
Db	2894	GTAATTGGACGGAGATGATGGGCGGACCCGCGAGCCGACCCCTCTCTC	2953
QY	61	ccgggctcgccagggcagggcgggagctccgcgacacacagcgggttctcagggc	120
Db	2954	CCGGGCTCGCGCAGGGCAGGGCGGGAGCTCCGCGCACCAACAGACGGCTTCTCAGGGC	3013
QY	121	gctttgctctctgtttttcccggttctgttttccctcttccggaaggttgtaa	180
Db	3014	GCTTTGCTCTCTGTTTTCCTCCGGTCTGTTTCTCCCTCTCCGGAAGGCTTGTCAA	3073
QY	181	ggggtagagaaagagacgcaaacacaaaagtggaaaacag	221
Db	3074	GGGCTAGGAGAAAGAGACGCAACACAAAAGTGGAACAG	3114
RESULT	6		
AX253452			
LOCUS	AX253452	7260 bp	DNA linear PAT 10-OCT-2001
DEFINITION	Sequence 3 from Patent WO0170810.		
ACCESSION	AX253452		
VERSION	AX253452.1	GI:16073979	
KEYWORDS	human.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 7260)		
AUTHORS	Schmitz, G. and Bodzioch, M.		
TITLE	Atp binding cassette transporter 1 (abcl) gene polymorphisms and uses thereof for the diagnosis and treatment of lipid, cardiovascular or inflammatory disorders		

100



```

repeat_region 41593..41835
/rpt_family="Alu"
STS
41657..41679
/gene="ABCA1"
/note="A009X28"
repeat_region 42320..42584
/rpt_family="Alu"
repeat_region 44349..44626
/rpt_family="Alu"
repeat_region 48473..48669
/rpt_family="Alu"
repeat_region 50547..50614
/rpt_family="Alu"
repeat_region 54285..54418
/rpt_family="Alu"
repeat_region 54849..54988
/rpt_family="Alu"
repeat_region 55029..55226
/rpt_family="Alu"
repeat_region 55297..55344
/rpt_family="Alu"
repeat_region 58501..58935
/note="LTR"
repeat_region 61246..61489
/rpt_family="HERV"
repeat_region 62812..63077
/rpt_family="Alu"
repeat_region 64082..64306
/rpt_family="Alu"
repeat_region 67123..67382
/rpt_family="Alu"
repeat_region 68499..68731
/rpt_family="Alu"
repeat_region 69481..69760
/rpt_family="Alu"
repeat_region 71998..72233
/rpt_family="Alu"
repeat_region 72244..72501
/rpt_family="Alu"

Query Match 100.0% Score 221; DB 9; Length 149034;
Best Local Similarity 100.0%; Pred. No. 7.9e-49;
Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gtaattgcagcagagtagtgaggcgccgagccgagcagcagccgagccctctctc 60
Db 1454 gtaattgcagcagagtagtgaggcgccgagccgagcagccgagccgagccctctc 1513

QY 61 ccgggtcgcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 120
Db 1514 ccgggtcgcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1573

QY 121 gctttgctctctgtttttcccccgttctgttttcccccctctccggaagcgttgc 180
Db 1574 gctttgctctctgtttttcccccgttctgttttcccccctctccggaagcgttgc 1633

QY 181 ggggtagagaagagacacacacacacacacacacacacacacacacacacacac 221
Db 1634 ggggtagagaagagacacacacacacacacacacacacacacacacacacacac 1674

RESULT 9
AC012230
LOCUS AC012230 175064 bp DNA linear HTG 22-APR-2000
DEFINITION Homo sapiens clone RP11-IM10, WORKING DRAFT SEQUENCE, 39 unordered
pieces.
ACCESSION AC012230
VERSION AC012230.3 GI:7637254
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 175064)  
Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
Homo sapiens, clone RP11-IM10  
Unpublished  
2 (bases 1 to 175064)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,  
Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhgalter,B.,  
Brown,A., Castle,A., Colangelo,M., Collins,S., Collumore,A.,  
Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,  
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D., Horton,L.,  
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,  
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,  
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,  
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,  
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,  
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,  
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

TITLE  
JOURNAL

Submitted (21-OCT-1999), Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Apr 22, 2000 this sequence version replaced gi:6454033.

COMMENT

All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence.submissions@genome.wi.mit.edu](mailto:sequence.submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L2510

Center clone name: L\_M10

----- Summary Statistics

Sequencing vector: M13; M7815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 117571 bases at least Q40

Consensus quality: 145749 bases at least Q30

Consensus quality: 160940 bases at least Q20

Insert size: 185000; agarose-fp

Insert size: 171264; sum-of-contigs

Quality coverage: 2.9 in Q20 bases; agarose-fp

Quality coverage: 3.2 in Q20 bases; sum-of-contigs

-----

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 39 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1003: contig of 1003 bp in length  
\* 1004 1103: gap of 100 bp  
\* 1104 2634: contig of 1531 bp in length  
\* 2635 2734: gap of 100 bp  
\* 2735 4415: contig of 1681 bp in length  
\* 4416 4515: gap of 100 bp  
\* 4516 5785: contig of 1270 bp in length  
\* 5786 5885: gap of 100 bp  
\* 5886 7879: contig of 1994 bp in length  
\* 7880 7979: gap of 100 bp  
\* 7980 9686: contig of 1707 bp in length  
\* 9687 9786: gap of 100 bp  
\* 9787 12253: contig of 2467 bp in length  
\* 12254 12353: gap of 100 bp  
\* 12354 15228: contig of 2875 bp in length  
\* 15229 15328: gap of 100 bp  
\* 15329 17200: contig of 1872 bp in length  
\* 17201 17300: gap of 100 bp









Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,  
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,  
Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,  
Zimmer,A. and Zody,M.

# TITLE

## JOURNAL

### COMMENT

Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jul 13, 2000 this sequence version replaced gi:6705871.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence.submissions@genome.wi.mit.edu](mailto:sequence.submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L2512

Center clone name: L\_N\_10

-----

\* NOTE: This record contains 73 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

1  
872 971: contig of 871 bp in length  
972 971: gap of 100 bp  
972 1834: contig of 863 bp in length  
1835 1934: gap of 100 bp  
1935 2804: contig of 870 bp in length  
2805 2904: gap of 100 bp  
2905 3745: contig of 841 bp in length  
3746 3845: gap of 100 bp  
3846 4696: contig of 851 bp in length  
4697 4796: gap of 100 bp  
4797 5640: contig of 844 bp in length  
5641 5740: gap of 100 bp  
5741 6540: contig of 800 bp in length  
6541 6640: gap of 100 bp  
6641 7509: contig of 869 bp in length  
7510 7609: gap of 100 bp  
7610 8479: contig of 870 bp in length  
8480 8579: gap of 100 bp  
8580 9430: contig of 851 bp in length  
9431 9530: gap of 100 bp  
9531 10376: contig of 846 bp in length  
10377 10476: gap of 100 bp  
10477 11322: contig of 846 bp in length  
11323 11422: gap of 100 bp  
11423 12302: contig of 880 bp in length  
12303 12402: gap of 100 bp  
12403 13280: contig of 878 bp in length  
13281 13380: gap of 100 bp  
13381 14241: contig of 861 bp in length  
14242 14341: gap of 100 bp  
14342 15196: contig of 855 bp in length  
15197 15296: gap of 100 bp  
15297 16123: contig of 827 bp in length  
16124 16223: gap of 100 bp  
16224 17072: contig of 849 bp in length  
17073 17172: gap of 100 bp  
17173 18041: contig of 869 bp in length  
18042 18141: gap of 100 bp  
18142 19009: contig of 868 bp in length  
19010 19109: gap of 100 bp  
19110 19966: contig of 857 bp in length

\* 19967 20066: gap of 100 bp  
20067 20921: contig of 855 bp in length  
20922 21021: gap of 100 bp  
21022 21865: contig of 844 bp in length  
21866 21965: gap of 100 bp  
21966 22832: contig of 867 bp in length  
22833 22932: gap of 100 bp  
22933 23780: contig of 848 bp in length  
23781 23880: gap of 100 bp  
23881 24733: contig of 853 bp in length  
24734 24833: gap of 100 bp  
24834 25670: contig of 837 bp in length  
25671 25770: gap of 100 bp  
25771 26621: contig of 851 bp in length  
26622 26721: gap of 100 bp  
26722 27576: contig of 855 bp in length  
27577 27676: gap of 100 bp  
27677 28532: contig of 856 bp in length  
28533 28632: gap of 100 bp  
28633 29492: contig of 860 bp in length  
29493 29592: gap of 100 bp  
29593 30455: contig of 863 bp in length  
30456 30555: gap of 100 bp  
30556 31410: contig of 855 bp in length  
31411 31510: gap of 100 bp  
31511 32368: contig of 858 bp in length  
32369 32468: gap of 100 bp  
32469 33312: contig of 844 bp in length  
33313 33412: gap of 100 bp  
33413 34268: contig of 856 bp in length  
34269 34368: gap of 100 bp  
34369 35204: contig of 836 bp in length  
35205 35304: gap of 100 bp  
35305 36156: contig of 852 bp in length  
36157 36256: gap of 100 bp  
36257 37128: contig of 872 bp in length  
37129 37228: gap of 100 bp  
37229 38083: contig of 855 bp in length  
38084 38183: gap of 100 bp  
38184 39031: contig of 848 bp in length  
39032 39131: gap of 100 bp  
39132 40006: contig of 875 bp in length  
40007 40106: gap of 100 bp  
40107 40967: contig of 861 bp in length  
40968 41067: gap of 100 bp  
41068 41913: contig of 846 bp in length  
41914 42013: gap of 100 bp  
42014 42824: contig of 811 bp in length  
42825 42924: gap of 100 bp  
42925 43776: contig of 852 bp in length  
43777 43876: gap of 100 bp  
43877 44752: contig of 876 bp in length  
44753 44852: gap of 100 bp  
44853 45724: contig of 872 bp in length  
45725 45824: gap of 100 bp  
45825 46643: contig of 819 bp in length  
46644 46743: gap of 100 bp  
46744 47599: contig of 856 bp in length  
47600 47699: gap of 100 bp  
47700 48551: contig of 852 bp in length  
48552 48651: gap of 100 bp  
48652 49485: contig of 834 bp in length  
49486 49585: gap of 100 bp  
49586 50440: contig of 855 bp in length  
50441 50540: gap of 100 bp  
50541 51404: contig of 864 bp in length  
51405 51504: gap of 100 bp  
51505 52372: contig of 868 bp in length  
52373 52472: gap of 100 bp  
52473 53328: contig of 856 bp in length  
53329 53428: gap of 100 bp  
53429 54268: contig of 840 bp in length  
54269 54368: gap of 100 bp

\* 54369 55229: contig of 861 bp in length  
 \* 55230 55329: gap of 100 bp  
 \* 55330 56197: contig of 868 bp in length  
 \* 56198 56297: gap of 100 bp  
 \* 56298 57163: contig of 866 bp in length  
 \* 57164 57263: gap of 100 bp  
 \* 57264 58130: contig of 867 bp in length  
 \* 58131 58230: gap of 100 bp  
 \* 58231 59082: contig of 852 bp in length  
 \* 59083 59182: gap of 100 bp  
 \* 59183 60032: contig of 838 bp in length

\* 59163 60020: contig of 636 bp in length  
\* 60021 60120: gap of 100 bp  
\* 50121 50093: contig of 863 bp in length

60121	8095: contig of 809 bp in length
60122	8095: contig of 100 bp
60984	61083: gap of 100 bp
61084	61935: contig of 852 bp in length
61936	62035: gap of 100 bp
62036	62866: contig of 831 bp in length
62867	62966: gap of 100 bp
62967	63827: contig of 961 bp in length
63827	63900: gap of 100 bp

```

* 63826 63927: gap of 100 bp
* 63928 64783: contig of 856 bp in length
* 64784 64883: gap of 100 bp
* 64884 65740: contig of 857 bp in length
* 65741 65840: gap of 100 bp
* 65841 66684: contig of 844 bp in length
* 66685 66784: gap of 100 bp
* 66785 67651: contig of 867 bp in length
* 67652 67751: gap of 100 bp

Query Match      99.5%  Score 220;  DB 2;  Length 69570;
Posterior Similarity 99.5%  Pred No. 1 4s-48;

```

best local similarity 99.5%, read no. 1,440, Matches 220; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 gtaattgcagcgagtgagtgaggccgagccgagccgagccgacccctctc 60  
|||||  
41565 GTAATTGCGAGCGAGGTGAGTGGGGCCGGGACCGCGAGCGGAGCGCCTTCTCTC 41624

y	61	ccgggctgcggcaggacaggcgaggagctccgcacacacagagcccggtlcttcagggc	120
b	41625	ccgggctgcggcaggacaggcgaggagctccgcacacacagagcccggttctcagggc	41684
y	121	gcttgcctcctggtttttcccccgtctggtttctcccttctccgaagccttgtaa	180
b	41685	gccttgcctcctggttttttcccccgttctggttttctcccttctccgaagccttgtaa	41744
y	181	ggggtaggagaaagacgcacaacacaaaagtggaaaaacag	221
b	41745	ggggtaggagaaagacgcacacacaaaagtggaaaaacag	41785

RESULT	13
K024328	
OCUCS	AK024328
DEFINITION	Homo sapiens cDNA FL114266 fis, clone PLACE1002437, highly similar to ATP-BINDING CASSETTE TRANSPORTER 1.
CCESSION	AK024328
VERSION	AK024328.1
EYWORDS	GI:10436685
SOURCE	oligo capping; fis (full insert sequence).
	Homo sapiens placenta cDNA to mRNA, clone_lib:PLACE1
	linear mRNA 1556 bp PRI 29-SEP-2000

ORGANISM	CLONE:PLACEL002437
Homo sapiens	
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.	
1 (sites)	
REFERENCE	
1. Iwase, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y.,	

TITLE  
 NEDO human cDNA sequencing project  
 unpublished (2000)  
 REFERENCES  
 1 (bases 1 to 1556)  
 2 (bases 1 and Otsuki, T.  
 Authors  
 Nishikawa, T., Nagai, K., K. Sugano, S., Takahashi-Fujii, A., Hara, H.,  
 Tanabe, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K.,  
 Ariita, M., Nabekura, T., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J.,  
 Wakamatsu, A., Nakamura, Y., Nagehara, K., Masuho, Y. and Oshima, A.



```
* 17779 18679: contig of 901 bp in length
* 18680 18779: gap of 100 bp
* 18780 19632: contig of 853 bp in length
* 19633 19732: gap of 100 bp
* 19733 20634: contig of 902 bp in length
* 20635 20734: gap of 100 bp
* 20735 21620: contig of 886 bp in length
* 21621 21720: gap of 100 bp
* 21721 22579: contig of 859 bp in length
* 22580 22679: gap of 100 bp
* 22680 23568: contig of 889 bp in length
* 23569 23668: gap of 100 bp
* 23669 24554: contig of 886 bp in length
* 24555 24654: gap of 100 bp
* 24655 25521: contig of 867 bp in length
* 25522 25621: gap of 100 bp
* 25622 26487: contig of 866 bp in length
* 26488 26587: gap of 100 bp
* 26588 27464: contig of 877 bp in length
* 27465 27564: gap of 100 bp
* 27565 28466: contig of 902 bp in length
* 28467 28566: gap of 100 bp
* 28567 29464: contig of 898 bp in length
* 29465 29564: gap of 100 bp
* 29565 30447: contig of 883 bp in length
* 30448 30547: gap of 100 bp
* 30548 31453: contig of 906 bp in length
* 31454 31553: gap of 100 bp
* 31554 32452: contig of 899 bp in length
* 32453 32552: gap of 100 bp
* 32553 33447: contig of 895 bp in length
* 33448 33547: gap of 100 bp
* 33548 34435: contig of 888 bp in length
* 34436 34535: gap of 100 bp
* 34536 35433: contig of 898 bp in length
* 35434 35533: gap of 100 bp
* 35534 36440: contig of 907 bp in length
* 36441 36540: gap of 100 bp
* 36541 37422: contig of 882 bp in length
* 37423 37522: gap of 100 bp
* 37523 38402: contig of 880 bp in length
* 38403 38502: gap of 100 bp
* 38503 39380: contig of 878 bp in length
* 39381 39480: gap of 100 bp
* 39481 40407: contig of 927 bp in length
* 40408 40507: gap of 100 bp
* 40508 41405: contig of 898 bp in length
* 41406 41505: gap of 100 bp
* 41506 42399: contig of 894 bp in length
* 42400 42499: gap of 100 bp
* 42500 43357: contig of 858 bp in length
* 43358 43457: gap of 100 bp
* 43458 44356: contig of 899 bp in length
* 44357 44456: gap of 100 bp
* 44457 45325: contig of 869 bp in length
* 45326 45425: gap of 100 bp
* 45426 46305: contig of 880 bp in length
* 46306 46405: gap of 100 bp
* 46406 47302: contig of 897 bp in length
* 47303 47402: gap of 100 bp
* 47403 48293: contig of 891 bp in length
* 48294 48393: gap of 100 bp
* 48394 49280: contig of 887 bp in length
* 49281 49380: gap of 100 bp
* 49381 50250: contig of 870 bp in length
* 50251 50350: gap of 100 bp
* 50351 51209: contig of 859 bp in length
* 51210 51309: gap of 100 bp
* 51310 52221: contig of 912 bp in length
* 52222 52321: gap of 100 bp
* 52322 53205: contig of 884 bp in length
* 53206 53305: gap of 100 bp
* 53306 54225: contig of 920 bp in length
```

```
* 54226 54325: gap of 100 bp
* 54326 55206: contig of 881 bp in length
* 55207 55306: gap of 100 bp
* 55307 56165: contig of 859 bp in length
* 56166 56265: gap of 100 bp
* 56266 57124: contig of 859 bp in length
* 57125 57224: gap of 100 bp
* 57225 58121: contig of 897 bp in length
* 58122 58221: gap of 100 bp
* 58222 59075: contig of 854 bp in length
* 59076 59175: gap of 100 bp
* 59176 60058: contig of 883 bp in length
* 60059 60158: gap of 100 bp
* 60159 61067: contig of 909 bp in length
* 61068 61167: gap of 100 bp
* 61168 62051: contig of 884 bp in length
* 62052 62151: gap of 100 bp
* 62152 63022: contig of 871 bp in length
* 63023 63122: gap of 100 bp
* 63123 64026: contig of 904 bp in length
* 64027 64126: gap of 100 bp
* 64127 65015: contig of 889 bp in length
* 65016 65115: gap of 100 bp
* 65116 65975: contig of 860 bp in length
* 65976 66075: gap of 100 bp
* 66076 66954: contig of 879 bp in length
* 66955 67054: gap of 100 bp
* 67055 67947: contig of 893 bp in length
* 67948 68047: gap of 100 bp
* 68048 68934: contig of 887 bp in length
* 68935 69034: gap of 100 bp
* 69035 69910: contig of 876 bp in length
* 69911 70010: gap of 100 bp

Query Match      97.2%; Score 214.8; DB 2; Length 90698;
Best Local Similarity 99.1%; Pred. NO. 3.5e-47;
Matches 216; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 attgcagcagagtgatggggccggagcccgagagccgagccgacctctctcccg 63
Db 87531 AATGCGAGCGAGAGTGGGGCCGGGACCGCGAGAGCCGCGAGAGCCGCTCTCCCG 87472

QY 64 ggctcgcgcagggcggggagctccgcgcacacacagagcggttctcaggcgct 123
Db 87471 GGCTCGCGCAGGCGAGGCGGGGAGCTCGCGCGCACCAACAGAGCGGTCTCAGGGCGCT 87412

QY 124 ttgctctctgtttttcccccgttctgttttctcccttctccggaaggcttgcagg 183
Db 87411 TTGCTCCTTGTGTTTTTCCCGGCTTCTGTTTTCTCCCTTCTCCGAAGGCTTGTCAAGG 87352

QY 184 gtaggagaaagacgacacacacacacacacacacacacacacacacacacacac 221
Db 87351 GTAGGAGAAAGAGACGCGCACACACACACACACACACACACACACACACACACAG 87314
```

Search completed: September 20, 2002, 03:14:51  
Job time: 14005 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 03:08:57 ; Search time 520.94 seconds  
(without alignments)  
728.372 Million cell updates/sec

Title: US-09-846-456-4  
Perfect score: 221  
Sequence: 1 gtaattgcgagcgagtgatg.....aacacaaaagtgaaacacag 221

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues 3472872

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_032802:\*  
1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1980.DAT:\*  
2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1981.DAT:\*  
3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1982.DAT:\*  
4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1983.DAT:\*  
5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1984.DAT:\*  
6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1985.DAT:\*  
7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1986.DAT:\*  
8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1987.DAT:\*  
9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1988.DAT:\*  
10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1989.DAT:\*  
11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1990.DAT:\*  
12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1991.DAT:\*  
13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1992.DAT:\*  
14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1993.DAT:\*  
15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1994.DAT:\*  
16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1995.DAT:\*  
17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1996.DAT:\*  
18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1997.DAT:\*  
19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1998.DAT:\*  
20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA1999.DAT:\*  
21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA2000.DAT:\*  
22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT:\*  
23: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA2001B.DAT:\*  
24: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	221	100.0	7260	22 AAD21326	Human ATP binding
2	221	100.0	7260	22 AAI70315	Human ATP binding
3	220.6	99.8	183999	22 AAF92831	Human ABC1 genomic
4	219.4	99.3	736	22 AAH07432	Human CDNA clone (
5	219.4	99.3	1556	22 AAH18606	Human CDNA sequenc
6	219	99.1	763	22 AAH04729	Human CDNA clone (
7	219	99.1	1750	22 AAH17451	Human CDNA sequenc
8	206	93.2	7281	22 AAK51683	Human polynucleoti
9	205.4	92.9	7086	22 ABA09200	Human ABCA1 homolo

10	205.4	92.9	7086	22 AAK52667	Human polynucleoti
11	205	92.8	9854	22 AAS06121	Human ABC1 DNA seq
12	197	89.1	227	21 AAC09615	Human secreted pro
13	197	89.1	10442	22 AAF24680	Nucleotide sequenc
14	197	89.1	10442	22 AAF24702	Nucleotide sequenc
15	190	86.0	10474	22 AAF24685	Nucleotide sequenc
16	190	86.0	10474	22 AAF24686	Nucleotide sequenc
17	190	86.0	10474	22 AAF24707	Nucleotide sequenc
18	190	86.0	10474	22 AAF24708	Nucleotide sequenc
19	92	41.6	446	22 AAS04035	Partial human ABC1
20	92	41.6	9741	22 AAS06120	Human ABC1 DNA seq
21	91	41.2	1643	22 AAF24681	Nucleotide sequenc
22	91	41.2	1643	22 AAF24703	Nucleotide sequenc
23	34.4	15.6	1704	22 AAS60816	Human cancer agent
24	34.2	15.5	10442	22 AAF24680	Nucleotide sequenc
25	34.2	15.5	10442	22 AAF24702	Nucleotide sequenc
26	34.2	15.5	10474	22 AAF24685	Nucleotide sequenc
27	34.2	15.5	10474	22 AAF24686	Nucleotide sequenc
28	34.2	15.5	10474	22 AAF24707	Nucleotide sequenc
29	34.2	15.5	10474	22 AAF24708	Nucleotide sequenc
30	33.6	15.2	117213	19 AAV62176	HSV-2 strain SB5 C
31	33.6	15.2	154746	24 AAD25519	Human herpesvirus
32	33.6	15.2	154746	24 AAD25519	Human herpesvirus
33	32.8	14.8	466	22 ABA12925	Human nervous syst
34	32.6	14.8	3250	20 AAX00726	Human aggregan deg
35	32.6	14.8	5530	21 AAK75435	Human OREF ORF990
36	32.6	14.8	29411	22 AAK76613	Human immune/haema
37	32.4	14.7	13001	22 AAK82979	Human immune/haema
38	32.2	14.6	669	23 AAS83553	DNA encoding novel
39	32.2	14.6	14332	22 AAS41688	Genomic sequence #
40	32.2	14.6	14332	22 AAK81810	Human immune/haema
41	32.2	14.6	14332	22 AAK84386	Human immune/haema
42	32.2	14.6	114955	20 AAX53491	Human adenosine A1
43	32.2	14.6	114955	20 AAX53491	Human adenosine A1
44	32	14.5	831	22 AAH06449	Human CDNA clone (
45	31.6	14.3	16831	22 AAK73455	Human immune/haema

## ALIGNMENTS

RESULT 1

AAD21326  
ID AAD21326 standard; DNA; 7260 BP.

XX AC AAD21326;

XX DT 28-JAN-2002 (first entry)

XX DE Human ATP binding cassette transporter 1 (ABC1) gene.

XX KW Human; ATP binding cassette transporter 1; ABC1; coronary heart disease; dermatological; atherosclerosis; cardiovascular; inflammatory disease; psoriasis; lipid disorder; antibacterial; septic shock; gene therapy; immunosuppressive; lupus erythematosus; rheumatoid arthritis; ds.

XX OS Homo sapiens.

XX PH Key Location/Qualifiers

XX FT CDS 321..7106

XX FT FT /\*tag= a /product= "Human ABC1 protein"

XX PN EPI136552-A1.

XX PD 26-SEP-2001.

XX PF 20-MAR-2000; 2000EP-0105820.

XX PR 20-MAR-2000; 2000EP-0105820.

XX PA (FARB ) BAYER AG.

PI Schmitz G, Bodzioch M;  
 XX WPI; 2001-640389/74.  
 DR P-PSDB; AAE13022.  
 XX  
 XX New adenosine triphosphate binding cassette transporter-1 gene  
 FT polymorphisms, useful for diagnosing and treating lipid disorders,  
 PT cardiovascular diseases and inflammatory diseases -  
 XX  
 XX Example 1; Fig 1; 48pp; English.  
 XX  
 CC The invention relates to four common polymorphisms in the gene encoding  
 CC ATP-binding cassette transporter-1 (ABCI). ABCI is associated with  
 CC decreased ApoA-I mediated efflux of cholesterol. The polymorphisms in  
 CC ABCI directly affects cellular lipid homeostasis, which is a key factor  
 CC in the atherogenetic processes. The ABCI polymorphisms are useful for  
 CC diagnosing and treating lipid disorders, cardiovascular diseases  
 CC (coronary heart disease, atherosclerosis) and inflammatory diseases  
 CC (psoriasis, lupus erythematosus). The identification of ABCI as a  
 CC transporter for interleukin-1beta (IL-1beta) identifies this gene as  
 CC a candidate for treatment of inflammatory diseases including rheumatoid  
 CC arthritis and septic shock. The present sequence is human ABCI gene.  
 XX  
 XX Sequence 7260 BP; 1834 A; 1765 C; 1905 G; 1756 T; 0 other;

Query Match 100.0%; Score 221; DB 22; Length 7260;  
 Best Local Similarity 100.0%; Pred. No. 8.7e-55;  
 Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 gtaattcgagcgagagtgagtgaggcgagccgagccgagcgagccgagcccttctc 60  
 Db |||||||  
 QY 8 gtaattcgagcgagagtgagtgaggcgagccgagccgagcgagccgagcccttctc 67  
 Db |||||||  
 QY 61 ccgggctgcgagcgagcgagcgagcgagccgagccgagccgagccgagccgagcc 120  
 Db |||||||  
 QY 68 ccgggctgcgagcgagcgagcgagcgagccgagccgagccgagccgagccgagcc 127  
 Db |||||||  
 QY 121 gctttgctctgtttttcccggttctgttttctctctctccgagcgttctca 180  
 Db |||||||  
 QY 128 gctttgctctgtttttcccggttctgttttctctctctccgagcgttctca 187  
 Db |||||||  
 QY 181 ggggtaggaagagagcgagcgagcgagcgagcgagcgagcgagcgagcgagcg 221  
 Db |||||||  
 QY 188 ggggtaggaagagagcgagcgagcgagcgagcgagcgagcgagcgagcgagcg 228  
 Db |||||||

RESULT 2  
 AAI70315  
 ID AAI70315 standard; cDNA; 7260 BP.  
 XX  
 AC AAI70315;  
 XX  
 XX 07-JAN-2002 (first entry)  
 DT  
 XX  
 DE Human ATP binding cassette transporter 1 (ABCI) cDNA.  
 XX  
 KW ATP binding cassette transporter 1; ABCI; human; lipid disorder;  
 KW cholesterol; cardiovascular disease; inflammatory disease;  
 KW antiinflammatory; antilipemic; antipsoriasis; dermatological;  
 KW Tangier disease; coronary heart disease; diagnosis; gene therapy;  
 KW polymorphism; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH 321..7106  
 FT /\*tag= a  
 FT CDS 501..7106  
 FT /\*tag= b  
 FT /\*tag= c  
 FT /\*note= "alternative open reading frame of AAI70314"  
 FT variation replace(976,A)  
 FT /\*tag= c

FT variation replace(1516,C)  
 FT /\*tag= d  
 FT variation replace(2969,G)  
 FT /\*tag= e  
 FT variation replace(3836,C)  
 FT /\*tag= f  
 XX EP1136554-A1.  
 XX  
 XX 26-SEP-2001.  
 XX  
 XX 24-MAR-2000; 2000EP-0106401.  
 XX  
 XX 24-MAR-2000; 2000EP-0106401.  
 XX  
 XX (FARB ) BAYER AG.  
 XX  
 XX Schmitz G, Bodzioch M;  
 XX  
 XX WPI; 2001-640389/74.  
 DR P-PSDB; AAM50228.  
 XX  
 XX New adenosine triphosphate binding cassette transporter gene  
 PT polymorphisms, useful for diagnosing and treating lipid disorders,  
 PT cardiovascular diseases and inflammatory diseases -  
 XX  
 XX Disclosure; Page 26-28; 41pp; English.  
 PS  
 XX  
 CC The present sequence is that of cDNA encoding the human adenosine  
 CC triphosphate (ATP) binding cassette transporter 1 (ABCI) protein  
 CC (see AAM50227). The sequence includes an extended open reading  
 CC frame (ORF) to that provided by the sequence in AAI70314, using  
 CC an alternative ATG codon as initiation codon and thereby adding an  
 CC extra 40 N-terminal amino acids to the encoded ABCI protein (see  
 CC AAM50228). The invention provides 4 common polymorphisms in the  
 CC ABCI gene. These were identified by sequencing the ABCI gene in  
 CC different Tangier kindreds. In the variant genes (numbering as in  
 CC AAI70314), G is changed to A at position 596, T is changed to C at  
 CC position 1136, A is changed to G at position 2589 or G is changed  
 CC to C at position 3456, or any combination of these. All of these  
 CC polymorphisms alter the amino acid sequence of ABCI and therefore  
 CC may affect its function. The 2 most common polymorphisms (G596A)  
 CC and A2589G) are both associated with a decreased in vitro ApoA-I  
 CC mediated efflux of cholesterol from mononuclear phagocytes, a  
 CC feature typical of Tangier disease. 3 Of the variants (G596A,  
 CC A2589G and G3456C) are significantly increased in a population of  
 CC men having low high density lipoprotein-cholesterol levels and  
 CC established coronary heart disease (CHD) relative to CHD-free  
 CC control subjects. The use of the provided ABCI polymorphisms for  
 CC the diagnosis and treatment of lipid disorders, cardiovascular  
 CC diseases, and inflammatory diseases (e.g. psoriasis, lupus  
 CC erythematosus) is claimed. Modulation of ABCI transcripts or  
 CC proteins by antisense or ribozyme technology or RNA decoys is also  
 CC claimed.  
 XX  
 XX Sequence 7260 BP; 1834 A; 1765 C; 1905 G; 1756 T; 0 other;

Query Match 100.0%; Score 221; DB 22; Length 7260;  
 Best Local Similarity 100.0%; Pred. No. 8.7e-55;  
 Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 gtaattcgagcgagagtgagtgaggcgagccgagccgagcgagccgagcccttctc 60  
 Db |||||||  
 QY 8 gtaattcgagcgagagtgagtgaggcgagccgagccgagcgagccgagcccttctc 67  
 Db |||||||  
 QY 61 ccgggctgcgagcgagcgagcgagcgagccgagccgagccgagccgagccgagcc 120  
 Db |||||||  
 QY 68 ccgggctgcgagcgagcgagcgagcgagccgagccgagccgagccgagccgagcc 127  
 Db |||||||  
 QY 121 gctttgctctgtttttcccggttctgttttctctctctccgagcgttctca 180  
 Db |||||||  
 QY 128 gctttgctctgtttttcccggttctgttttctctctctccgagcgttctca 187  
 Db |||||||



QY 181 ggggtaggagaaagagacgcaacacaaagtgaataacag 221  
|||||  
Db 188 ggggtaggagaaagagacgcaacacaaagtgaataacag 228  
|||||

RESULT 3  
AAF92831  
ID AAF92831 standard; DNA; 183999 BP.

AC AAF92831;

XX 17-MAY-2001 (first entry)

XX Human ABC1 genomic DNA.

XX High density lipoprotein-cholesterol; HDL-C; cardiovascular; ABC1; ds.

XX Homo sapiens.

XX WO200115676-A2.

XX 08-MAR-2001.

XX 01-SEP-2000; 2000WO-IB01492.

XX 01-SEP-1999; 99US-0151977.

XX 15-MAR-2000; 2000US-0526193.

XX 23-JUN-2000; 2000US-0213958.

XX (UYBR-) UNIV BRITISH COLUMBIA.

XX (XENO-) XENON GENETICS INC.

XX Hayden MR, Brooks-Wilson AR, Pimstone SN, Clee SM;

XX WPI; 2001-244356/25.

XX Treating a lower than normal high density lipoprotein-cholesterol  
(HDL-C) level, a higher than normal triglyceride level, or a  
PT cardiovascular disease, by administering a compound that modulates LXR-  
PT or RXR-mediated transcriptional activity -

XX Claim 8; Fig 1; 317pp; English.

XX The present invention relates to a method for treating a patient  
CC diagnosed as having a lower than normal high density

CC lipoprotein-cholesterol (HDL-C) level, a higher than normal

CC triglyceride level, or a cardiovascular disease, involving

CC administering a compound that modulates LXR- or RXR-mediated

CC transcriptional activity or ABC1 expression or activity.

CC The LXR gene product may be used in an assay to identify

CC compounds useful for the treatment of a disease or condition selected a

CC lower than normal HDL cholesterol level, a higher than normal

CC triglyceride level, and a cardiovascular disease.

XX Sequence 183999 BP; 49549 A; 37944 C; 41170 G; 54950 T; 386 other;

Query Match 99.8%; Score 220.6; DB 22; Length 183999;

Best Local Similarity 99.5%; Pred. No. 3.1e-54;

Matches 220; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 gtaattcgagcagagagtgagtgaggccgagccgagagccgagccaccttcttc 60  
|||||

Db 28790 gtaattcgagcagagagtgagtgaggccgagccgagagccgagccaccttcttc 28849  
|||||

QY 61 ccgggctgcggcagggcgagggcgaggagctcccgccacacagagccggttctcagggc 120  
|||||

Db 28850 ccgggctgcggcagggcgagggcgaggagctcccgccacacagagccggttctcagggc 28909  
|||||

QY 121 gctttgctcttcttttttcccggttcttcttttcccttctccggaagggttctca 180  
|||||

Db 28910 gctttgctcttctttttcccggttcttcttttcccttctccggaagggttctca 28969  
|||||

QY 181 ggggtaggagaaagagacgcaacacaaagtgaataacag 221  
|||||  
Db 28970 ggggtaggagaaagagacgcaacacaaagtgaataacag 29010  
|||||

RESULT 4  
AAH07432  
ID AAH07432 standard; cDNA; 736 BP.

XX AAH07432;

XX 26-JUN-2001 (first entry)

XX Human cDNA clone (5'-primer) SEQ ID NO:4267.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602

XX full-length cDNAs defined in the specification, and for the detection

XX and/or diagnosis of the abnormality of the proteins encoded by the

XX full-length cDNAs -

XX Claim 1; SEQ ID 4267; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602

XX full-length cDNAs defined in the specification. Where a primer set

XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary

XX to the complementary strand of a polynucleotide which comprises one of

XX the 5602 nucleotide sequences defined in the specification, where the

XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination

XX of an oligonucleotide comprising a sequence complementary to the

XX complementary strand of a polynucleotide which comprises a 5'-end

XX sequence and an oligonucleotide comprising a sequence complementary to a

XX polynucleotide which comprises a 3'-end sequence, where the

XX oligonucleotide comprises at least 15 nucleotides, and the combination of

XX the 5'-end sequence/3'-end sequence is selected from those defined in

XX the specification. The primer sets can be used in antisense therapy and

XX in gene therapy. The primers are useful for synthesizing polynucleotides,

XX particularly full-length cDNAs. The primers are also useful for the

XX detection and/or diagnosis of the abnormality of the proteins encoded by

XX the full-length cDNAs. The primers allow obtaining of the full-length

XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and

XX AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to

XX AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632

XX represent oligonucleotides, all of which are used in the exemplification

XX of the present invention.

XX Sequence 736 BP; 163 A; 199 C; 199 G; 170 T; 5 other;

Query Match 99.3%; Score 219.4; DB 22; Length 736;



The present invention describes primer sets for synthesising 5602  
 full-length cDNAs defined in the specification. Where a primer set  
 comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 to the complementary strand of a polynucleotide which comprises one of  
 the 5602 nucleotide sequences defined in the specification, where the  
 oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 of an oligonucleotide comprising a sequence complementary to the  
 complementary strand of a polynucleotide which comprises a 5'-end  
 sequence and an oligonucleotide comprising a sequence complementary to a  
 polynucleotide which comprises a 3'-end sequence, where the  
 oligonucleotide comprises at least 15 nucleotides and the combination of  
 the 5'-end sequence/3'-end sequence is selected from those defined in  
 the specification. The primer sets can be used in antisense therapy and  
 in gene therapy. The primers are useful for synthesising polynucleotides  
 particularly full-length cDNAs. The primers are also useful for the  
 detection and/or diagnosis of the abnormality of the proteins encoded by  
 the full-length cDNAs. The primers allow obtaining of the full-length  
 cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 AAH13633 to AAH18742 represent human cDNA sequences; AAB32446 to  
 AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 represent oligonucleotides, all of which are used in the exemplification  
 of the present invention.

Sequence 763 BP; 137 A; 205 C; 260 G; 158 T; 3 other;

(HELI-) HELIX RES INST.

Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
WPI; 2001-318749/34.

Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs -

claim 8; SEQ ID 16905; 2537pp + CD ROM; English.

The present invention describes primer sets for synthesising 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesising polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

Sequence 1750 BP; 291 A; 489 C; 586 G; 384 T; 0 other;

Query Match	99.1%;	Score 219;	DB 22;	Length 1750;
Best Local Similarity	100.0%;	Pred. No. 2.2e-54;		
Matches 219;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	3	aattgcagcagagtgagtgaggccgggacccgcagagccgagccgaccccttctctccc	62
Db	1	aattgcagcagagtgagtgaggccgggacccgcagagccgagccgaccccttctctccc	60
QY	63	gggctgcggcagggcagggcgggagctccgcgcaccacacagagccggttctcagggcgc	122
Db	61	gggctgcggcagggcagggcgggagctccgcgcaccacacagagccggttctcagggcgc	120
QY	123	tttgcctctgtttttcccggttctgttttctcccttctccggaaggcttgtcaagg	182
Db	121	tttgcctctgtttttcccggttctgttttctcccttctccggaaggcttgtcaagg	180
QY	193	ggtaggacaagagacgcaaacacaaaagtggaaaacag	221
Db	191	ggtaggacaagagacgcaaacacaaaagtggaaaacag	219

RESULT	8
AAK51683	
ID	AAK51683 standard; cDNA; 7281 BP.
XX	
XX	AAK51683;
XX	
XX	06-NOV-2001 (first entry)
XX	
XX	Human polynucleotide SEQ ID NO 228.
DE	
XX	Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW	





PD	03-MAY-2001.	
XX		
XX	26-OCT-2000; 2000WO-EF10886.	
XX		
XX	26-OCT-1999; 99EP-0402668.	
PR	01-MAR-2000; 2000US-0186260.	
XX		
XX	(AVET ) AVENTIS PHARMA SA.	
XX		
XX	Denefle P, Rosier-Montus M, Arnould-Reguigne I, Prades C, Naudin L;	
PI	Lemoline C, Duverger N, Jaye M, Searfoss GH, Remailley A, Brewer HB;	
PI	Dean M;	
XX		
XX	WPI; 2001-316327/33.	
XX	P-PSDB; AAU02176.	
XX		
XX	New human ABC1 nucleic acids and polypeptides for treating	
PT	atherosclerosis, malaria and diabetes -	
PT		
PS	Claim 1; Page 209-213; 368pp; English.	
XX		
CC	The sequence represents the coding sequence #2 of human ABC1. The	
CC	nucleic acid sequence, primers and probes derived from the ABC1 sequence,	
CC	and polypeptides and vectors are useful for the prevention of	
CC	atherosclerosis, in a subject affected by a dysfunction in the reverse	
CC	transport of cholesterol. The polypeptide encoded by the ABC1 gene is	
CC	useful for screening for an active ingredient for the prevention or	
CC	treatment of a disease resulting from dysfunction in the reverse	
CC	transport of cholesterol. The nucleic acids and polypeptides are also	
CC	useful for treating and preventing cardiovascular and neurological	
CC	pathologies, and other diseases e.g. Tangier disease, lecithin-	
CC	cholesterol (LCAT) deficiency, malaria and diabetes.	
XX		
XX	Sequence 9854 BP; 2665 A; 2219 C; 2334 G; 2635 T; 1 other:	

Query Match	92.8%;	Score 205;	DB 22;	Length 9854;
Best Local Similarity	100.0%;	Pred. No. 4.6e-50;		
Matches 205;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	17	gtgagtgggcccgggaccgcagagccgagccgaccccttctccgggctgcggcaggg	76	
Db	1	gtgagtgggcccgggaccgcagagccgagccgaccccttctccgggctgcggcaggg	60	
QY	77	caggsgggggagctccgcgcaccaacagaccggtttctcagggcgtttgctccctgttt	136	
Db	61	caggsgggggagctccgcgcaccaacagaccggtttctcagggcgtttgctccctgttt	120	
QY	137	tttcccggttctgtttttccctctctccgaagccttgcacaggggtaggagaagag	196	
Db	121	tttcccggttctgtttttccctctctccgaagccttgcacaggggtaggagaagag	180	
QY	197	acgcaaacacaaaagtggaaaacag	221	
Db	181	acgcaaacacaaaagtggaaaacag	205	

RESULT 12  
AAC09615  
ID AAC09615 standard; CDNA; 227 BP.

```

XX 06-SEP-2000.
PD
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
XX
PR 26-FEB-1999; 99US-0122487.
XX
XX (GEST ) GENSET.
XX
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX
XX WPI; 2000-500381/45.
XX
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX Claim 1; SEQ ID 13690; 71pp + CD-ROM; English.
XX
XX The present sequence is one of a large number of 5' ESTs derived from
XX mRNAs encoding secreted proteins. No ORF has yet been conclusively
XX identified within the present sequence. The 5' ESTs were prepared from
XX total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
XX sequences usually correspond mainly to the 3' untranslated region (UTR)
XX of the mRNA because they are not well obtained from oligo-dT primed cDNA
XX libraries. Such ESTs are not often suited for isolating cDNA sequences
XX derived from the 5' ends of mRNAs and even in those cases where longer
XX cDNA sequences have been obtained, the full 5' UTR is rarely included.
XX 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
XX used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
XX in diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX They are used to obtain upstream regulatory sequences and to design
XX expression and secretion vectors.
XX
XX Sequence 227 BP; 44 A; 65 C; 73 G; 45 T; 0 other;

Query Match      89.1%; Score 197; DB 21; Length 227;
Best Local Similarity 100.0%; Pred. No. 3,1e-48;
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gtaattcgagcagagatgtagtggcgccggagccagagccagagccaccctctctc 60
Db 31 gtaattcgagcagagatgtagtggcgccggagccagagccagagccaccctctc 90
QY 61 ccgggctcgcgagcagggcagggcgagctccgcgcaccaacagagccggttctcagggc 120
Db 91 ccgggctcgcgagcagggcagggcgagctccgcgcaccaacagagccggttctcagggc 150
QY 121 gctttgctcttgtttttcccggttctgttttctcccttctccggaagcgttgtcaa 180
Db 151 gctttgctcttgtttttcccggttctgttttctcccttctccggaagcgttgtcaa 210
QY 181 ggggtaggagaaagaga 197
Db 211 ggggtaggagaaagaga 227

RESULT 13
AAAF24680
ID AAF24680 standard; DNA; 10442 BP.
XX
XX AAF24680;
AC
XX
XX
XX 20-APR-2001 (first entry)
XX
XX Nucleotide sequence of a human ABC1 polypeptide.
XX
XX Human; adenosine triphosphate binding cassette protein 1; ABC1;
KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
KW chromosome 9q22-q31; heart disease; hypercholesterolemia;
KW atherosclerosis; cholesterol transport; ss.
XX

```

```

OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH CDS 291..7076
FT /*tag= a
FT /*product= "ABCl polypeptide"
XX
XX WO200078972-A2.
XX
XX 28-DEC-2000.
XX
XX 16-JUN-2000; 2000WO-US16765.
XX
XX 18-JUN-1999; 99US-0140264.
XX
XX 14-SEP-1999; 99US-0153872.
XX
XX 19-NOV-1999; 99US-0166573.
XX
XX (CVTH-) CV THERAPEUTICS INC.
XX
XX Lawn RM, Wade D, Garvin M;
XX
XX WPI; 2001-137812/14.
XX
XX Adenosine triphosphate (ATP) binding cassette (ABC) polynucleotide,
XX useful for the development of agents for the treatment of heart disease
XX and other disorders associated with hypercholesterolemia and
XX atherosclerosis -
XX
XX Disclosure; Page 122-128; 215pp; English.
XX
XX The present sequence encodes a human adenosine triphosphate (ATP)
XX binding cassette protein (ABC) 1 polypeptide. ABC1 resides in cell
XX membranes and utilises ATP hydrolysis to transport a wide variety of
XX substrates across the plasma membrane. ABC1 is a pivotal protein in
XX the apolipoprotein-mediated mobilisation of intracellular cholesterol
XX stores. ABC1 is defective in Tangier disease, a genetic disorder
XX characterised by abnormal HDL-cholesterol metabolism. The ABC1 gene is
XX localised to chromosome 9q22-9q31. The ABC1 genes and proteins are
XX useful for developing pharmaceutical agents for the treatment of heart
XX disease and other disorders associated with hypercholesterolemia and
XX atherosclerosis. The genes are useful for developing screening assays to
XX screen for compounds that regulate the expression of genes associated
XX with cholesterol transport. The genes and proteins are also useful for
XX with cholesterol transport. The genes and proteins are also useful for
XX other disorders associated with hypercholesterolemia.
XX
XX Sequence 10442 BP; 2898 A; 2297 C; 2408 G; 2835 T; 4 other;

Query Match 89.1%; Score 197; DB 22; Length 10442;
Best Local Similarity 100.0%; Pred. No. 1e-47;
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 ggcgggaccgcagagccgagccgaccttctctcccggtgcgagggcagggcg 84
Db 1 ggcgggaccgcagagccgagccgaccttctctcccggtgcgagggcagggcg 60
QY 85 ggagctccgcgacacagcagcggtctctcagggcggttgcctcttcttttcccg 144
Db 61 ggagctccgcgacacagcagcggtctctcagggcggttgcctcttcttttcccg 120
QY 145 gtctcttttctcccttctccgaggttctcaggggttagagaaagagacgcaaac 204
Db 121 gtctcttttctcccttctccgaggttctcaggggttagagaaagagacgcaaac 180
QY 205 acaaaagtggaaaacag 221
Db 181 acaaaagtggaaaacag 197

RESULT 14
AAF24702
ID AAF24702 standard; DNA; 10442 BP.

```

```

XX
XX AAF24702;
XX
XX 20-APR-2001 (first entry)
XX
XX Nucleotide sequence of a human ABC1 polypeptide.
XX
XX Human; adenosine triphosphate binding cassette protein 1; ABC1;
XX apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
XX chromosome 9q22-9q31; heart disease; hypercholesterolemia;
XX atherosclerosis; cholesterol transport; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH CDS 291..7076
FT /*tag= a
FT /*product= "ABCl polypeptide"
XX
XX WO200078971-A2.
XX
XX 28-DEC-2000.
XX
XX 16-JUN-2000; 2000WO-US16591.
XX
XX 18-JUN-1999; 99US-0140264.
XX
XX 14-SEP-1999; 99US-0153872.
XX
XX 19-NOV-1999; 99US-0166573.
XX
XX (CVTH-) CV THERAPEUTICS INC.
XX (UNIW ) UNIV WASHINGTON.
XX
XX Lawn RM, Wade D, Oram JF, Garvin M;
XX
XX WPI; 2001-137811/14.
XX
XX P-PSDB; AAB31365.
XX
XX Adenosine triphosphate (ATP) binding cassette protein (ABC) 1
XX polynucleotides and polypeptides, useful for treatment of heart disease
XX and other disorders associated with hypercholesterolemia and
XX atherosclerosis -
XX
XX Claim 3; Page 117-123; 211pp; English.
XX
XX The present sequence encodes a human adenosine triphosphate (ATP)
XX binding cassette protein (ABC) 1 polypeptide. ABC1 resides in cell
XX membranes and utilises ATP hydrolysis to transport a wide variety of
XX substrates across the plasma membrane. ABC1 is a pivotal protein in
XX the apolipoprotein-mediated mobilisation of intracellular cholesterol
XX stores. ABC1 is defective in Tangier disease, a genetic disorder
XX characterised by abnormal HDL-cholesterol metabolism. The ABC1 gene is
XX localised to chromosome 9q22-9q31. The ABC1 genes and proteins are
XX useful for developing pharmaceutical agents for the treatment of heart
XX disease and other disorders associated with hypercholesterolemia and
XX atherosclerosis. The genes are useful for developing screening assays to
XX screen for compounds that regulate the expression of genes associated
XX with cholesterol transport. The genes and proteins are also useful for
XX with cholesterol transport. The genes and proteins are also useful for
XX other disorders associated with hypercholesterolemia.
XX
XX Sequence 10442 BP; 2898 A; 2297 C; 2408 G; 2835 T; 4 other;

Query Match 89.1%; Score 197; DB 22; Length 10442;
Best Local Similarity 100.0%; Pred. No. 1e-47;
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 ggcgggaccgcagagccgagccgaccttctctcccggtgcgagggcagggcg 84
Db 1 ggcgggaccgcagagccgagccgaccttctctcccggtgcgagggcagggcg 60
QY 85 ggagctccgcgacacagcagcggttctcagggcggttgcctcttcttttcccg 144

```











US-08-327-832-4





```

RESULT          9
US-07-885-972A-3/c
; Sequence 3, Application US/07885972A
; Patent No. 5460950
; GENERAL INFORMATION:
; APPLICANT: Barr, Philip J.
; APPLICANT: Brake, Anthony J.
; APPLICANT: Kaufman, Rnadal J.
; APPLICANT: Tekamp-Olson, Patricia
; APPLICANT: Wasley, Louise
; APPLICANT: Wong, Polly A.

```

RESULT 10  
US-08-745-880-3/c  
: Sequence 3, Application US/08745880  
: Patent No. 5965425  
: GENERAL INFORMATION:  
: APPLICANT: Barr, Philip J.  
: APPLICANT: Brake, Anthony J.  
: APPLICANT: Kaufman, Rnadal J.  
: APPLICANT: Tekamp-Olson, Patricia





```

RESULT 12
US-09-335-409-1
; Sequence 1, Application US/09335409
; Patent No. 6121029
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOPHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/335,409
; CURRENT FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 68750
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
US-09-335-409-1

Query Match 13.4%; Score 29.6; DB 3; Length 68750;
Best Local Similarity 61.8%; Pred. No. 35;
Matches 47; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

```

QY 15 g<sub>af</sub>g<sub>at</sub>g<sub>gg</sub>g<sub>cc</sub>g<sub>gg</sub>g<sub>acc</sub>g<sub>c</sub>g<sub>a</sub>g<sub>ac</sub>g<sub>cc</sub>g<sub>a</sub>g<sub>cc</sub>g<sub>ac</sub>cc<sub>t</sub>t<sub>c</sub>c<sub>c</sub>g<sub>g</sub>g<sub>at</sub>g<sub>cg</sub>g<sub>c</sub>g<sub>a</sub>g 74  
 |||||  
 Db 18310 g<sub>gt</sub>g<sub>ag</sub>g<sub>t</sub>g<sub>ccc</sub>t<sub>g</sub>a<sub>c</sub>g<sub>ac</sub>g<sub>cc</sub>g<sub>c</sub>g<sub>c</sub>g<sub>at</sub>g<sub>ct</sub>g<sub>ct</sub>g<sub>cg</sub>g<sub>c</sub>g<sub>ac</sub>g<sub>cc</sub>g<sub>c</sub>g<sub>a</sub>g 18369  
 |||||  
 QY 75 g<sub>c</sub>g<sub>g</sub>g<sub>g</sub>g<sub>g</sub>g<sub>g</sub>g<sub>at</sub>g<sub>ct</sub> 90  
 |||||  
 Db 18370 c<sub>at</sub>c<sub>g</sub>c<sub>g</sub>c<sub>g</sub>c<sub>g</sub>g<sub>at</sub>g<sub>ct</sub> 18385  
 |||||

```

RESULT 13
US-09-568-102-1
; Sequence 1, Application US/09568102
; Patent No. 6346404
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOPHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,102
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 68750
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
US-09-568-102-1

```

Query Match	13.4%	Score 29.6;	DB 4;	Length 68750;
Best Local Similarity	61.8%;	Pred. No. 35;		
Matches 47;	Conservative 0;	Mismatches 29;	Indels 0;	Gaps 0;
QY 15	gagtgagtgggcggggacccgagagccgagccgctctctccggcgctcggcgag 74			
Db 18310	gggtgagcgtggcgtgagcaacgccgcgctcgacggctgctctcgggcgagcgcgcgag 18369			
QY 75	ggcaggggcggggagct 90			

```

Db 18370 gcatcgcgaggtgct 18385
|||||
RESULT 14
US-09-567-969-1
; Sequence 1, Application US/09567969
; Patent NO. 6355457
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/567,969
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 68750
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
US-09-567-969-1

```

	Query Match	13.4%	Score 29.6;	DB 4;	Length 68750;
	Best Local Similarity	61.8%;	Pred. No. 35;		
	Matches 47;	Conservative 0;	Mismatches 29;	Indels 0;	Gaps 0;
QY	15	gagtgagtggggcgagccgcgagcgcgagccgacctctctccgggctcggcgag	74		
Db	18310	gggtgagcgtggcgtgagcaacgccgcgcgtcgacgctgctctcgggcgagcgcgag	18369		
QY	75	ggcagggcggggagct	90		
Db	18370	cqatcgacgaagtct	18385		

```

RESULT 15
US-09-568-480-1
; Sequence 1, Application US/09568480
; Patent No. 6355458
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,480
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 68750
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
US-09-568-480-1

```

Query Match	13.4%	Score	29.6;	DB	4;	Length	68750;
Best Local Similarity	61.8%;	Pred.	No. 35;				
Matches	47;	Conservative	0;	Mismatches	29;	Indels	0;
						Gaps	0;



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 03:14:51 ; Search time 5225.75 Seconds  
(without alignments)  
636.716 Million cell updates/sec

Title: US-09-846-456-5

Perfect score: 159

Sequence: 1 ttaatgaccagccacggcg.....cttcagaagaagacaaaca 159

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_htg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vi.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vi.\*

30: em\_htg\_hum.\*

31: em\_htg\_inv.\*

32: em\_htg\_other.\*

33: em\_htgo\_inv.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	ID	Description
-----					

1	159	100.0	159	6	AX351033	Sequence
2	159	100.0	357	6	AX351030	Sequence
3	155.8	98.0	10442	6	AX351030	Sequence
4	155.8	98.0	10442	6	AX060713	Sequence
5	155.8	98.0	10442	9	AX060892	Sequence
6	155.8	98.0	10474	6	AF285167	Homo sapi
7	155.8	98.0	10474	6	AX060719	Sequence
8	155.8	98.0	10474	6	AX060721	Sequence
9	155.8	98.0	10474	6	AX060898	Sequence
10	155.8	98.0	10474	6	AX060900	Sequence
11	153.2	96.4	183999	9	AF275948	Homo sapi
12	143.4	90.2	697	9	AX092589	Sequence
13	142.2	89.4	200	9	AF258627	Homo sapi
14	142.2	89.4	446	6	AF258624	Homo sapi
15	142.2	89.4	446	6	AX127764	Sequence
16	142.2	89.4	480	6	AX139751	Sequence
17	142.2	89.4	7260	6	HS252277	Homo sapi
18	142.2	89.4	9741	6	AX253452	Sequence
19	142.2	89.4	9741	6	AX127830	Sequence
20	142.2	89.4	9741	6	AX139817	Sequence
21	142.2	89.4	9854	6	AX351038	Sequence
22	142.2	89.4	9854	6	AX127831	Sequence
23	142.2	89.4	129608	9	AX139818	Sequence
24	142.2	89.4	175064	2	AL353685	Human DNA
25	142.2	89.4	201144	2	AC012230	Homo sapi
26	140.6	88.4	1556	9	AF287262	Homo sapi
27	136.2	85.7	298	9	AK024328	Homo sapi
28	135.2	85.0	7860	6	AB037924	Homo sapi
29	135.2	85.0	7862	6	AB092594	Sequence
30	130.2	81.9	90698	2	AX135712	Sequence
31	72.8	45.8	207659	2	AC021345	Homo sapi
32	72.8	45.8	278572	2	AC091466	Mus muscu
33	67	42.1	7878	10	AF287263	Mus muscu
34	64.4	40.5	6786	9	MMABCL	Mus muscu
35	54	34.0	7074	5	AB055982	Mus muscu
36	37	23.3	110000	2	AF362377	Homo sapi
37	35.8	22.5	61524	9	LMFLCHR36_08	Continuation (9 of
38	35.6	22.4	52376	2	AL353725	Human DNA
39	35.4	22.3	37	6	AC102919	Mus muscu
40	35	22.0	22881	10	AX092843	Sequence
41	35	22.0	23024	6	AF287142	Mus muscu
42	35	22.0	185825	10	AX080494	Sequence
43	35	22.0	200143	2	AC087114	Mus muscu
44	34.6	21.8	6704	9	AC073805	Mus muscu
45	34.6	21.8	6768	6	AF328787	Homo sapi
					AX320362	Sequence

## ALIGNMENTS

RESULT	1	AX351033	Sequence	5 from Patent WO0183746.	159 bp	DNA	linear	PAT 06-FEB-2002
LOCUS	AX351033	Sequence	5 from Patent WO0183746.	159 bp	DNA	linear	PAT 06-FEB-2002	
DEFINITION	AX351033	Sequence	5 from Patent WO0183746.	159 bp	DNA	linear	PAT 06-FEB-2002	
ACCESSION	AX351033	Sequence	5 from Patent WO0183746.	159 bp	DNA	linear	PAT 06-FEB-2002	
VERSION	AX351033.1	GI:18616389						
KEYWORDS	human.							
SOURCE	human.							
ORGANISM	Homo sapiens							
REFERENCE	1 (sites)							
AUTHORS	Rosier-Montus,M.F., Prades,C., Lemoine,C., Naudin,L., Deneffe,P., Brewer,B., Duverger,N., Renaley,A. and Santamarina-Fojo,S.							
TITLE	Regulatory nucleic acid sequences of the abcl gene							
JOURNAL	Patent: WO 0183746-A 5 08-NOV-2001;							
FEATURES	Aventis Pharma S.A. (FR)							
Source	Location/Qualifiers							
	1..159							
	/organism="Homo sapiens"							
	/db_xref="taxon:9606"							
BASE COUNT	32 a	47 c	48 g	32 t				
ORIGIN								

```
Query Match 100.0%; Score 159; DB 6; Length 159;
Best Local Similarity 100.0%; Pred. No. 1.7e-33;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttaatgaccagccacggggcgctccctctctcagctctgagcctgcccgtcctccagggtccc 60
DB 1 TTAATGACCAGCCACGGGGCGTCCCTGCTGTCAGCTGTGGCCGCTGCTTCCAGGGCTCCC 60

QY 61 gagccacacgctggcgctgctgaggaacatgcatgttgctgagctgaggttg 120
DB 61 GAGCCACACGCTGGCGTGTGCTGAGGAACATGSCATGTGGCTCAGCTGAGGTG 120

QY 121 ctgctgtggaagaacctcactttcagaagaagacaaaca 159
DB 121 CTGCTGTGGAAGAACCTCACTTTTCAGAAGAAGACAAACA 159

RESULT 2
AX351030 AX351030 357 bp DNA linear PAT 06-FEB-2002
LOCUS Sequence 2 from Patent WO0183746.
DEFINITION AX351030
ACCESSION AX351030
VERSION AX351030.1 GI:18616386
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (sites)
AUTHORS Rosier-Montus M.F., Prades C., Lemoine C., Naudin L., Denefle P.,
Brewer B., Duverger N., Remaley A. and Santamarina-Fojo S.
TITLE Regulatory nucleic acid sequences of the abcl gene
JOURNAL Patent: WO 0183746-A 2 08-NOV-2001;
Aventis Pharma S.A. (FR)
FEATURES
source
1. .357
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 64 a 78 c 116 g 99 t
ORIGIN

Query Match 100.0%; Score 159; DB 6; Length 357;
Best Local Similarity 100.0%; Pred. No. 1.6e-33;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ttaatgaccagccacggggcgctccctctctcagctctgagcctgcccgtcctccagggtccc 60
DB 100 TTAATGACCAGCCACGGGGCGTCCCTGCTGTCAGCTGTGGCCGCTGCTTCCAGGGCTCCC 159

QY 61 gagccacacgctggcgctgctgaggaacatgcatgttgctgagctgaggttg 120
DB 160 GAGCCACACGCTGGCGTGTGCTGAGGAACATGSCATGTGGCTCAGCTGAGGTG 219

QY 121 ctgctgtggaagaacctcactttcagaagaagacaaaca 159
DB 220 CTGCTGTGGAAGAACCTCACTTTTCAGAAGAAGACAAACA 258

RESULT 3
AX060713 AX060713 10442 bp DNA linear PAT 22-JAN-2001
LOCUS Sequence 1 from Patent WO0078972.
DEFINITION AX060713
ACCESSION AX060713
VERSION AX060713.1 GI:12406103
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 10442)
AUTHORS Lawn, R.M., Wade, D., Oram, J.F. and Garvin, M.
TITLE Atp binding cassette transporter protein abcl polypeptides
JOURNAL Patent: WO 0078971-A 1 28-DEC-2000;
CV THERAPEUTICS, INC. (US)
FEATURES
source
1. .10442
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2898 a 2297 c 2408 g 2835 t
ORIGIN

Query Match 98.0%; Score 155.8; DB 6; Length 10442;
Best Local Similarity 98.7%; Pred. No. 1.1e-32;
Matches 157; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ttaatgaccagccacggggcgctccctctctcagctctgagcctgcccgtcctccagggtccc 60
DB 198 TTAATGACCAGCCACGGGGCGTCCCTGCTGTCAGCTGTGGCCGCTGCTTCCAGGGCTCCC 257

QY 61 gagccacacgctggcgctgctgaggaacatgcatgttgctgagctgaggttg 120
DB 258 GAGCCACACGCTGGCGTGTGCTGAGGAACATGSCATGTGGCTCAGCTGAGGTG 317

QY 121 ctgctgtggaagaacctcactttcagaagaagacaaaca 159
DB 318 CTGCTGTGGAAGAACCTCACTTTTCAGAAGAAGACAAACA 356

RESULT 4
AX060892 AX060892 10442 bp DNA linear PAT 22-JAN-2001
LOCUS Sequence 1 from Patent WO0078971.
DEFINITION AX060892
ACCESSION AX060892
VERSION AX060892.1 GI:12406270
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 10442)
AUTHORS Lawn, R.M., Wade, D., Oram, J.F. and Garvin, M.
TITLE Atp binding cassette transporter protein abcl polypeptides
JOURNAL Patent: WO 0078971-A 1 28-DEC-2000;
CV THERAPEUTICS, INC. (US)
FEATURES
source
1. .10442
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2898 a 2297 c 2408 g 2835 t
ORIGIN

Query Match 98.0%; Score 155.8; DB 6; Length 10442;
Best Local Similarity 98.7%; Pred. No. 1.1e-32;
Matches 157; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ttaatgaccagccacggggcgctccctctctcagctctgagcctgcccgtcctccagggtccc 60
DB 198 TTAATGACCAGCCACGGGGCGTCCCTGCTGTCAGCTGTGGCCGCTGCTTCCAGGGCTCCC 257

QY 61 gagccacacgctggcgctgctgaggaacatgcatgttgctgagctgaggttg 120
DB 258 GAGCCACACGCTGGCGTGTGCTGAGGAACATGSCATGTGGCTCAGCTGAGGTG 317

QY 121 ctgctgtggaagaacctcactttcagaagaagacaaaca 159
DB 318 CTGCTGTGGAAGAACCTCACTTTTCAGAAGAAGACAAACA 356

RESULT 5
AF285167 AF285167 10442 bp mRNA linear PRI 09-AUG-2000
LOCUS
```













Db 212 CTGCTGTGGAGAACCTCCTCTTTCAGAGAGACAAACA 250

Search completed: September 20, 2002, 03:15:13  
Job time: 14027 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 01:23:14 ; Search time 3895 seconds  
(without alignments)  
765.809 Million cell updates/sec

Title: US-09-846-456-4

Perfect score: 221

Sequence: 1 gtaattgcgagcgagtgatg.....aacacaaaagtggaacacag 221

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	219.4	99.3	736	9	AU135588
2	219	99.1	763	9	AU121731
3	204	92.3	232	10	Z44377
4	178.2	80.6	998	10	BG678861
5	67.4	30.5	547	12	AZ769996
6	67.4	30.5	619	9	BB657864
7	65.4	29.6	625	9	BB665939
8	57.8	26.5	218	9	BB594197
9	57.8	26.2	259	9	BB594265
10	55.2	25.0	272	9	BB568993
11	50.4	22.8	276	9	BB570397
12	47.4	21.4	535	10	BG384217
13	39.2	17.7	925	12	CNS0091P
14	38.8	17.6	811	12	BH470300
15	38.4	17.4	910	12	CNS0060N
16	37.4	16.9	879	12	CNS0200G
17	37.2	16.8	311	9	BB374442

18	37.2	16.8	594	10	C76305
19	37.2	16.8	1051	12	AG127909
20	37	16.7	700	10	BF237183
21	37	16.7	981	12	AG131276
22	36.8	16.7	315	12	CNS01677
23	36.8	16.7	1015	12	AG137074
24	36.6	16.6	313	9	BB318923
25	36.6	16.5	904	10	BI101161
26	36.2	16.4	353	9	AI850822
27	36.2	16.4	429	9	AW489757
28	36.2	16.4	457	9	AW490154
29	36.2	16.4	466	10	BE948350
30	36.2	16.4	476	9	BB698363
31	36.2	16.4	489	10	BE943674
32	36.2	16.4	489	10	BE943841
33	36.2	16.4	493	10	BE947468
34	36.2	16.4	557	9	BB696895
35	36.2	16.4	650	9	AV334840
36	36.2	16.4	656	9	BB547934
37	36	16.3	1158	12	AG073892
38	35.6	16.1	320	9	BB317508
39	35.6	16.1	400	10	BB692826
40	35.6	16.1	409	10	BI082717
41	35.6	16.1	411	9	AA098273
42	35.6	16.1	411	9	BB790023
43	35.6	16.1	442	10	BF224565
44	35.6	16.1	442	10	BF319828
45	35.6	16.1	446	9	AA562016

## ALIGNMENTS

RESULT 1

AU135588

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AU135588 736 bp mRNA linear EST 24-OCT-2000  
AU135588 PLACE1 Homo sapiens cDNA clone PLACE1002437 5', mRNA  
sequence.

AU135588.1 GI:10996127

human.

EST.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 736)

Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,

Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and

Isogai,T.

HRI human cDNA project

Unpublished (2000)

Contact: Takao Isogai

Genomics Laboratory

Helix Research Institute

1532-3 Yana, Kisarazu, Chiba 292-0812, Japan

Tel: 81-438-52-3951

Fax: 81-438-52-3952

Email: genomics@hri.co.jp

HRI human cDNA project; 5' - & 3' - end one pass sequencing: Helix

Research Institute; cDNA library construction: Department of

Virology, Institute of Medical Science, University of Tokyo, and

Helix Research Institute.

Location/Qualifiers

1..736

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="PLACE1002437"

/clone\_lib="PLACE1"

/tissue\_type="placenta"

/note="Vector: pME18SFL3"

163 a 199 c 199 g 170 t

5 others

BASE COUNT

ORIGIN

















Ov 9 a

34; Conservative 80; Mismatches 66; Indels 0; Gaps 0;

QY 9 gagcagagtgtggggccgggacccgcagagccgagccgaccttctctcccggtg 68



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 03:09:16 ; Search time 520.94 Seconds  
(without alignments)  
524.032 Million cell updates/sec

Title: US-09-846-456-5

Perfect score: 159

Sequence: 1 ttaatgaccagccacggcg.....ctttcagaagaagacaaaca 159

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_032802.\*

- 1: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/NA1980.DAT.\*
- 2: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/NA1981.DAT.\*
- 3: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/NA1982.DAT.\*
- 4: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/NA1983.DAT.\*
- 5: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/NA1984.DAT.\*
- 6: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/NA1985.DAT.\*
- 7: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/NA1986.DAT.\*
- 8: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/NA1987.DAT.\*
- 9: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/NA1988.DAT.\*
- 10: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/NA1989.DAT.\*
- 11: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/NA1990.DAT.\*
- 12: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/NA1991.DAT.\*
- 13: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/NA1992.DAT.\*
- 14: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/NA1993.DAT.\*
- 15: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/NA1994.DAT.\*
- 16: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/NA1995.DAT.\*
- 17: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/NA1996.DAT.\*
- 18: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/NA1997.DAT.\*
- 19: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/NA1998.DAT.\*
- 20: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/NA1999.DAT.\*
- 21: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/NA2000.DAT.\*
- 22: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT.\*
- 23: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/NA2001B.DAT.\*
- 24: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	155.8	98.0	10442	22	AAF24680 Nucleotide sequenc
2	155.8	98.0	10442	22	AAF24702 Nucleotide sequenc
3	155.8	98.0	10474	22	AAF24685 Nucleotide sequenc
4	155.8	98.0	10474	22	AAF24686 Nucleotide sequenc
5	155.8	98.0	10474	22	AAF24707 Nucleotide sequenc
6	155.8	98.0	10474	22	AAF24708 Nucleotide sequenc
7	153.2	96.4	18399	22	AAF52831 Human ABC1 genomic
8	142.2	89.4	446	22	AA04035 Partial human ABC1
9	142.2	89.4	7086	22	ABA09200 Human ABCA1 homolo

10	142.2	89.4	7086	22	AAK52667 Human polynucleoti
11	142.2	89.4	7260	22	AA021326 Human ATP binding
12	142.2	89.4	7260	22	AAI70315 Human ATP binding
13	142.2	89.4	7281	22	AAK51683 Human polynucleoti
14	142.2	89.4	9741	22	AA061120 Human ABC1 DNA seq
15	142.2	89.4	9854	22	AA061121 Human ABC1 DNA seq
16	142.2	89.4	10545	21	AAK69132 Human ABC1 gene ex
17	140.6	88.4	736	22	AAH07432 Human cDNA clone (
18	140.6	88.4	1556	22	AAH18606 Human cDNA sequenc
19	135.2	85.0	7857	21	AAK69388 Human ABC1 cholest
20	135.2	85.0	7860	22	AAK69386 Human ABC1 cholest
21	135.2	85.0	7860	22	AAK69387 Human ABC1 cholest
22	135.2	85.0	7861	21	AAK69385 Human ABC1 cholest
23	135.2	85.0	7864	21	AAK69386 Human ABC1 cholest
24	135.2	85.0	7864	21	AAK69387 Human ABC1 cholest
25	135.2	85.0	7864	21	AAK69388 Human ABC1 cholest
26	135.2	85.0	7864	21	AAK69389 Human ABC1 cholest
27	35.8	22.5	50885	22	AAK70336 Human immune/haema
28	35.4	22.3	37	22	AAK93084 ABC1 polymorphism
29	35	22.0	298	21	AAK06182 Human secreted pro
30	35	22.0	23024	22	AAK25499 Nucleotide sequenc
31	34.6	21.8	534	21	AAK75311 Human ORFX ORF866
32	34.6	21.8	1133	22	AAK05589 Human secreted pro
33	33.2	20.9	5669	22	AA050870 Human PD-ATP-bind
34	33.2	20.9	6522	22	AA080706 Human PD-ATP-bind
35	31.4	19.7	6607	22	AAK54812 Nucleotide sequenc
36	31.4	19.7	8217	22	AAK07164 Human reproductive
37	30.4	19.1	485	21	AAK01004 Human secreted pro
38	30.4	19.1	532	21	AAK01005 Human secreted pro
39	30.4	19.1	2043	18	AAK79627 Human Doc2-beta ge
40	30.2	19.0	1093	22	ABA08699 Human ESRP1 protei
41	30.2	19.0	1727	21	AAK66225 Human cDNA encodin
42	30.2	19.0	2143	22	AAI59934 Human polynucleoti
43	30.2	19.0	2644	22	AAK77695 Human Traf4 bindin
44	30.2	19.0	2668	22	AAK44707 Novel protein kina
45	30.2	19.0	2853	21	AAK21857 Human breast and o

## ALIGNMENTS

RESULT 1

AAF24680

ID AAF24680 standard; DNA; 10442 BP.

AC AAF24680;

XX 20-APR-2001 (first entry)

DT Nucleotide sequence of a human ABC1 polypeptide.

DE Human; adenosine triphosphate binding cassette protein 1; ABC1;

KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;

KW chromosome 9q22-q31; heart disease; hypercholesterolemia;

KW atherosclerosis; cholesterol transport; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 291..7076

XX FT /\*tag= a

XX FT /product= "ABC1 polypeptide"

XX PN WO200078972-A2.

XX PD 28-DEC-2000.

XX PF 16-JUN-2000; 2000WO-US16765.

XX PR 18-JUN-1999; 99US-0140264.

XX PR 14-SEP-1999; 99US-0153872.

XX PR 19-NOV-1999; 99US-0166573.



XX OS Homo sapiens.  
 XX FH Key Location/Qualifiers  
 XX FT CDS 323..7108  
 XX FT /\*tag= a  
 XX FT /product= "defective ABC1 polypeptide"  
 XX PN WO200078972-A2.  
 XX PD 28-DEC-2000.  
 XX PF 16-JUN-2000; 200WO-US16765.  
 XX PR 18-JUN-1999; 99US-0140264.  
 XX PR 14-SEP-1999; 99US-0153872.  
 XX PR 19-NOV-1999; 99US-0166573.  
 XX XX (CVTH-) CV THERAPEUTICS INC.  
 XX PA Lawn RM, Wade D, Garvin M;  
 XX PI WPI; 2001-137812/14.  
 XX DR  
 XX XX Adenosine triphosphate (ATP) binding cassette (ABC) polynucleotide,  
 PT useful for the development of agents for the treatment of heart disease  
 PT and other disorders associated with hypercholesterolemia and  
 PT atherosclerosis -  
 XX PA (CVTH-) CV THERAPEUTICS INC.  
 XX XX Lawn RM, Wade D, Garvin M;  
 XX PI WPI; 2001-137812/14.  
 XX DR  
 XX XX The present sequence encodes a human adenosine triphosphate (ATP)  
 CC binding cassette protein (ABC) 1 polypeptide, and is isolated from  
 CC a tangier disease patient. ABC1 resides in cell membranes and utilises  
 CC ATP hydrolysis to transport a wide variety of substrates across the  
 CC plasma membrane. ABC1 is a pivotal protein in the apolipoprotein-mediated  
 CC mobilisation of intracellular cholesterol stores. ABC1 is defective in  
 CC Tangier disease, a genetic disorder characterised by abnormal  
 CC HDL-cholesterol metabolism. The ABC1 gene is localised to chromosome  
 CC 9q22-9q31. The ABC1 genes and proteins are useful for developing  
 CC pharmaceutical agents for the treatment of heart disease and other  
 CC disorders associated with hypercholesterolemia and atherosclerosis. The  
 CC genes are useful for developing screening assays to screen for compounds  
 CC that regulate the expression of genes associated with cholesterol  
 CC transport. The genes and proteins are also useful for are also useful  
 CC as diagnostic indicators of cardiovascular disease and other disorders  
 CC associated with hypercholesterolemia.  
 XX PA Sequence 10474 BP; 2906 A; 2305 C; 2416 G; 2843 T; 4 other;  
 XX SQ  
 Query Match 98.0%; Score 155.8; DB 22; Length 10474;  
 Best Local Similarity 98.7%; Pred. No. 1.7e-36;  
 Matches 157; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 ttaatgaccagcagcggtccctgctgctgagctctgagctgctgagctgctcc 60  
 Db 230 ttaatgaccagcagcggtccctgctgctgagctctgagctgctgagctgctcc 289  
 QY 61 gagccacacgctggcgctgctgagtgaggaacatggttgcctcagctgaggttg 120  
 Db 290 gagccacacgctggcgctgctgagtgaggaacatggttgcctcagctgaggttg 349  
 QY 121 ctgctgtggaagaacctcactttcagaagaagacaaca 159  
 Db 350 ctgctgtggaagaacctcactttcagaagaagacaaca 388  
 RESULT 4  
 AAF24686  
 ID AAF24686 standard; DNA; 10474 BP.  
 XX  
 AC AAF24686;

XX 20-APR-2001 (first entry)  
 XX DT Nucleotide sequence of ABC1 polypeptide from Tangier disease patient.  
 XX DE  
 XX KW Human; adenosine triphosphate binding cassette protein 1; ABC1;  
 KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;  
 KW chromosome 9q22-9q31; heart disease; hypercholesterolemia;  
 KW atherosclerosis; cholesterol transport; ss.  
 XX OS Homo sapiens.  
 XX FH Key Location/Qualifiers  
 XX FT CDS 323..7108  
 XX FT /\*tag= a  
 XX FT /product= "defective ABC1 polypeptide"  
 XX PN WO200078972-A2.  
 XX PD 28-DEC-2000.  
 XX PF 16-JUN-2000; 200WO-US16765.  
 XX PR 18-JUN-1999; 99US-0140264.  
 XX PR 14-SEP-1999; 99US-0153872.  
 XX PR 19-NOV-1999; 99US-0166573.  
 XX XX (CVTH-) CV THERAPEUTICS INC.  
 XX PA Lawn RM, Wade D, Garvin M;  
 XX PI WPI; 2001-137812/14.  
 XX DR  
 XX XX Adenosine triphosphate (ATP) binding cassette (ABC) polynucleotide,  
 PT useful for the development of agents for the treatment of heart disease  
 PT and other disorders associated with hypercholesterolemia and  
 PT atherosclerosis -  
 XX PA (CVTH-) CV THERAPEUTICS INC.  
 XX XX Lawn RM, Wade D, Garvin M;  
 XX PI WPI; 2001-137812/14.  
 XX DR  
 XX XX The present sequence encodes a human adenosine triphosphate (ATP)  
 CC binding cassette protein (ABC) 1 polypeptide, and is isolated from  
 CC a tangier disease patient. ABC1 resides in cell membranes and utilises  
 CC ATP hydrolysis to transport a wide variety of substrates across the  
 CC plasma membrane. ABC1 is a pivotal protein in the apolipoprotein-mediated  
 CC mobilisation of intracellular cholesterol stores. ABC1 is defective in  
 CC Tangier disease, a genetic disorder characterised by abnormal  
 CC HDL-cholesterol metabolism. The ABC1 gene is localised to chromosome  
 CC 9q22-9q31. The ABC1 genes and proteins are useful for developing  
 CC pharmaceutical agents for the treatment of heart disease and other  
 CC disorders associated with hypercholesterolemia and atherosclerosis. The  
 CC genes are useful for developing screening assays to screen for compounds  
 CC that regulate the expression of genes associated with cholesterol  
 CC transport. The genes and proteins are also useful for are also useful  
 CC as diagnostic indicators of cardiovascular disease and other disorders  
 CC associated with hypercholesterolemia.  
 XX PA Sequence 10474 BP; 2907 A; 2304 C; 2415 G; 2844 T; 4 other;  
 XX SQ  
 Query Match 98.0%; Score 155.8; DB 22; Length 10474;  
 Best Local Similarity 98.7%; Pred. No. 1.7e-36;  
 Matches 157; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 ttaatgaccagcagcggtccctgctgctgagctctgagctgctgagctgctcc 60  
 Db 230 ttaatgaccagcagcggtccctgctgctgagctctgagctgctgagctgctcc 289  
 QY 61 gagccacacgctggcgctgctgagtgaggaacatggttgcctcagctgaggttg 120  
 Db 290 gagccacacgctggcgctgctgagtgaggaacatggttgcctcagctgaggttg 349  
 QY 121 ctgctgtggaagaacctcactttcagaagaagacaaca 159





CC genes are useful for developing screening assays to screen for compounds  
 CC that regulate the expression of genes associated with cholesterol  
 CC transport. The genes and proteins are also useful for are also useful  
 CC as diagnostic indicators of cardiovascular disease and other disorders  
 CC associated with hypercholesterolemia.

XX SQ Sequence 10474 BP; 2907 A; 2304 C; 2415 G; 2844 T; 4 other;

Query Match 98.0%; Score 155.8; DB 22; Length 10474;  
 Best Local Similarity 98.7%; Pred. No. 1.7e-36;  
 Matches 157; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ttaatgaccagccacggcgctccctgctgtcagctctggcgcgtgctccagggctccc 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 230 ttaatgaccagccacggcgctccctgctgtgagctctggcgcgtgctccagggctccc 289  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 61 gagccacacgctggcgctgctggcgtgaggaacatggcctgctcagctgaggttg 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 290 gagccacacgctggcgctgctggcgtgaggaacatggcctgctcagctgaggttg 349  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 121 ctgctgtggaagaacctcactttcagaagaagacaaaca 159  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 350 ctgctgtggaagaacctcactttcagaagaagacaaaca 388  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 7  
 ID AAF92831  
 XX AAF92831 standard; DNA; 183999 BP.  
 AC AAF92831;  
 DT 17-MAY-2001 (first entry)  
 DE Human ABC1 genomic DNA.  
 KW High density lipoprotein-cholesterol; HDL-C; cardiovascular; ABC1; ds.  
 OS Homo sapiens.  
 PN WO200115676-A2.  
 XX 08-MAR-2001.  
 PF 01-SEP-2000; 2000WO-IB01492.  
 PR 01-SEP-1999; 99US-0151977.  
 PR 15-MAR-2000; 2000US-0526193.  
 PR 23-JUN-2000; 2000US-0213958.  
 XX (UYBR-) UNIV BRITISH COLUMBIA.  
 PA (XENO-) XENON GENETICS INC.  
 XX Hayden MR, Brooks-Wilson AR, Pimstone SN, Clee SM;  
 WPI; 2001-244356/25.

Treating a lower than normal high density lipoprotein-cholesterol  
 (HDL-C) level, a higher than normal triglyceride level, or a  
 cardiovascular disease, by administering a compound that modulates LXR-  
 or RXR-mediated transcriptional activity -  
 Claim 8; Fig 1; 317pp; English.  
 The present invention relates to a method for treating a patient  
 diagnosed as having a lower than normal high density  
 lipoprotein-cholesterol (HDL-C) level, a higher than normal  
 triglyceride level, or a cardiovascular disease, involving  
 administering a compound that modulates LXR- or RXR-mediated  
 transcriptional activity or ABC1 expression or activity.  
 The LXR gene product may be used in an assay to identify  
 compounds useful for the treatment of a disease or condition selected a  
 lower than normal HDL cholesterol level, a higher than normal

CC triglyceride level, and a cardiovascular disease.

XX SQ Sequence 183999 BP; 49549 A; 37944 C; 41170 G; 54950 T; 386 other;

Query Match 96.4%; Score 153.2; DB 22; Length 183999;  
 Best Local Similarity 97.5%; Pred. No. 2.1e-35;  
 Matches 155; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ttaatgaccagccacggcgctccctgctgtcagctctggcgcgtgctccagggctccc 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 53229 ttaatgaccagccacggcgctccctgctgtgagctctggcgcgtgctccagggctccc 53288  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 61 gagccacacgctggcgctgctggcgtgaggaacatggcctgctcagctgaggttg 120  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 53289 gagccacacgctggcgctgctggcgtgaggaacatggcctgctcagctgaggttg 53348  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 QY 121 ctgctgtggaagaacctcactttcagaagaagacaaaca 159  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 53349 ctgctgtggaagaacctcactttcagaagaagacaaaca 53387  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 8  
 ID AAS04035  
 XX AAS04035 standard; cDNA; 446 BP.  
 AC AAS04035;  
 DT 12-SEP-2001 (first entry)  
 DE Partial human ABC1 cDNA sequence.  
 KW Human; ABC1 gene; atherosclerosis; reverse transport; cholesterol;  
 KW cardiovascular; neurological; Tangier disease; LCAT deficiency;  
 KW lecithin-cholesterol acetyltransferase; malaria; diabetes; ss.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT CDS 185..438  
 FT /\*tag= a  
 FT /product= "Human ABC1 protein, amino acids 1 to 60"  
 PN WO200130848-A2.  
 XX 03-MAY-2001.  
 PF 26-OCT-2000; 2000WO-EP10886.  
 PR 26-OCT-1999; 99EP-0402668.  
 PR 01-MAR-2000; 2000US-0186260.  
 XX (AVET) AVENTIS PHARMA SA.  
 XX Denefle P, Rosier-Montus M, Arnould-Requigne I, Prades C, Naudin L;  
 PI Lemoine C, Duverger N, Jaye M, Searfoss GH, Remaley A, Brewer HB;  
 PI Dean M;  
 XX WPI; 2001-316327/33.  
 DR P-PSDB; AAU02176.

PT New human ABC1 nucleic acids and polypeptides for treating  
 atherosclerosis, malaria and diabetes -  
 Example 2; Page 167; 368pp; English.  
 The sequence represents the partial coding sequence of human ABC1,  
 CC which encodes amino acids 1-60 of the human ABC1 protein. The nucleic  
 CC acid sequence, primers and probes derived from the ABC1 sequence, and  
 CC polypeptides and vectors are useful for the prevention of  
 CC atherosclerosis, in a subject affected by a dysfunction in the reverse  
 CC transport of cholesterol. The polypeptide encoded by the ABC1 gene is  
 CC useful for screening for an active ingredient for the prevention or





AC AAI70315;  
 XX  
 DT 07-JAN-2002 (first entry)  
 XX  
 DE Human ATP binding cassette transporter 1 (ABCI) cDNA.  
 XX  
 KW ATP binding cassette transporter 1; ABC1; human; lipid disorder;  
 KW cholesterol; cardiovascular disease; inflammatory disease;  
 KW antiinflammatory; antilipemic; antiporiatic; dermatological;  
 KW Tangier disease; coronary heart disease; diagnosis; gene therapy;  
 KW polymorphism; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 321..7106  
 FT /\*tag= a  
 FT CDS 501..7106  
 FT /\*tag= b  
 FT /\*note= "alternative open reading frame of AAI70314"  
 FT variation replace(976,A)  
 FT /\*tag= c  
 FT variation replace(1516,C)  
 FT /\*tag= d  
 FT variation replace(2969,G)  
 FT /\*tag= e  
 FT variation replace(3836,C)  
 FT /\*tag= f  
 XX  
 XX EPI136554-A1.  
 XX  
 XX 26-SEP-2001.  
 XX  
 XX 24-MAR-2000; 2000EP-0106401.  
 XX  
 XX 24-MAR-2000; 2000EP-0106401.  
 XX  
 XX (FARB ) BAYER AG.  
 XX  
 XX Schmitz G, Bodzioch M;  
 XX WPI; 2001-640389/74.  
 XX P-PSDB; AAM50228.  
 XX  
 XX New adenosine triphosphate binding cassette transporter gene  
 XX polymorphisms, useful for diagnosing and treating lipid disorders,  
 XX cardiovascular diseases and inflammatory diseases  
 XX  
 XX Disclosure; Page 26-28; 41pp; English.  
 XX  
 XX The present sequence is that of cDNA encoding the human adenosine  
 XX triphosphate (ATP) binding cassette transporter 1 (ABCI) protein  
 XX (see AAM50227). The sequence includes an extended open reading  
 XX frame (ORF) to that provided by the sequence in AAI70314, using  
 XX an alternative ATG codon as initiation codon and thereby adding an  
 XX extra 40 N-terminal amino acids to the encoded ABC1 protein (see  
 XX AAM50228). The invention provides 4 common polymorphisms in the  
 XX ABC1 gene. These were identified by sequencing the ABC1 gene in  
 XX different Tangier kindreds. In the variant genes (numbering as in  
 XX AAI70314), G is changed to A at position 596, T is changed to C at  
 XX position 1136, A is changed to G at position 2589 or G is changed  
 XX to C at position 3456, or any combination of these. All of these  
 XX polymorphisms alter the amino acid sequence of ABC1 and therefore  
 XX may affect its function. The 2 most common polymorphisms (G596A  
 XX and A2589G) are both associated with a decreased in vitro ApoA-I  
 XX mediated efflux of cholesterol from mononuclear phagocytes, a  
 XX feature typical of Tangier disease. 3 Of the variants (G596A,  
 XX A2589G and G3456C) are significantly increased in a population of  
 XX men having low high density lipoprotein-cholesterol levels and  
 XX established coronary heart disease (CHD) relative to CHD-free  
 XX control subjects. The use of the provided ABC1 polymorphisms for  
 XX the diagnosis and treatment of lipid disorders, cardiovascular  
 XX diseases, and inflammatory diseases (e.g. psoriasis, lupus

CC erythematoses) is claimed. Modulation of ABC1 transcripts or  
 CC proteins by antisense or ribozyme technology or RNA decoys is also  
 CC claimed.  
 XX  
 SQ Sequence 7260 BP; 1834 A; 1765 C; 1905 G; 1756 T; 0 other;  
 Query Match 89.4%; Score 142.2; DB 22; Length 7260;  
 Best Local Similarity 97.5%; Pred. No. 1.7e-32;  
 Matches 155; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
 QY 1 ttaatgaccagccacggcggtccctgctgtgcagctctggtgcgcgtgcctccagggtcc 60  
 |||||  
 DB 229 ttaatgaccagccac-gggtccctgctgtgagctctggtgcgcgtgcctccagggtcc 287  
 QY 61 gagccacacgctggcggtgctgggtgaggaacatggcatgttgccctcagctgaggttg 120  
 |||||  
 DB 288 gagccacacgctgggggtgctgggtgaggaacatggcttctggtgcctcagctgaggttg 347  
 QY 121 ctgctgtggaagaacctcactttcagaagaagacaaaca 159  
 |||||  
 DB 348 ctgctgtggaagaacctcactttcagaagaagacaaaca 386  
 RESULT 13  
 AAK51683  
 ID AAK51683 standard; cDNA: 7281 BP.  
 XX  
 AC AAK51683;  
 XX  
 XX 06-NOV-2001 (first entry)  
 XX  
 DE Human polynucleotide SEQ ID NO 228.  
 XX  
 DE Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorder; arthritis; inflammation; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200157190-A2.  
 XX  
 XX 09-AUG-2001.  
 XX  
 XX 05-FEB-2001; 2001WO-US04098.  
 XX  
 XX 03-FEB-2000; 2000US-0496914.  
 XX 27-APR-2000; 2000US-0560875.  
 XX 20-JUN-2000; 2000US-0598075.  
 XX 19-JUL-2000; 2000US-0620325.  
 XX 01-SEP-2000; 2000US-0654936.  
 XX 15-SEP-2000; 2000US-0663561.  
 XX 20-OCT-2000; 2000US-0693325.  
 XX 30-NOV-2000; 2000US-0728422.  
 XX (HYSE-) HYSEQ INC.  
 XX  
 XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;  
 XX Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
 XX Xue AJ, Yang Y, Wejhrman T, Goodrich R;  
 XX  
 XX WPI; 2001-476283/51.  
 XX P-PSDB; AAM78550.  
 XX  
 XX Nucleic acids encoding polypeptides with cytokine-like activities,  
 XX useful in diagnosis and gene therapy -  
 XX  
 XX Claim 1; Page 1086-1096; 6221pp; English.  
 XX  
 XX The invention relates to polynucleotides (AAK51456-AAK53435) and the  
 XX encoded polypeptides (AAM78323-AAAM80302) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce



XX New human ABC1 nucleic acids and polypeptides for treating  
PT atherosclerosis, malaria and diabetes -  
XX  
XX Claim 1; Page 209-213; 368pp; English.  
XX  
CC The sequence represents the coding sequence #2 of human ABC1. The  
CC nucleic acid sequence, primers and probes derived from the ABC1 sequence,  
CC and polypeptides and vectors are useful for the prevention of  
CC atherosclerosis, in a subject affected by a dysfunction in the reverse  
CC transport of cholesterol. The polypeptide encoded by the ABC1 gene is  
CC useful for screening for an active ingredient for the prevention or  
CC treatment of a disease resulting from dysfunction in the reverse  
CC transport of cholesterol. The nucleic acids and polypeptides are also  
CC useful for treating and preventing cardiovascular and neurological  
CC pathologies, and other diseases e.g. Tangier disease, lecithin-  
CC cholesterol (LCAT) deficiency, malaria and diabetes.  
XX  
SQ Sequence 9854 BP; 2665 A; 2219 C; 2334 G; 2635 T; 1 other;

Query Match 89.4%; Score 142.2; DB 22; Length 9854;  
Best Local Similarity 97.5%; Pred. NO. 1.8e-32;  
Matches 155; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
Qy 1 ttaatgaccagccagggcgctccctgctgcagctctgcccgtctgccttcagggtccc 60  
Db 206 ttaatgaccagccac-ggcgccctgctgagctctgcccgtctgccttcagggtccc 264  
Qy 61 gagccacacgctgggctgctgctgagggagacatggcatgttggcctcagctgaggttg 120  
Db 265 gagccacacgctggggtgctgctgagggagacatggttggcctcagctgaggttg 324  
Qy 121 ctgctgtggaagaacctcactttcagaagaagacaaaca 159  
Db 325 ctgctgtggaagaacctcactttcagaagaagacaaaca 363

Search completed: September 20, 2002, 03:09:36  
Job time: 10645 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 03:21:58 ; Search time 110.16 Seconds  
(without alignments)  
354.536 Million cell updates/sec

Title: US-09-846-456-5  
Perfect score: 159  
Sequence: 1 ttaatgaccagccagggcg.....cttcagaagaagacaaaca 159

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_NA:\*

1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq.\*

2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq.\*

3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq.\*

4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq.\*

5: /cgn2\_6/ptodata/2/ina/PCRU.COMB.seq.\*

6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	18.9	4771	2	US-08-866-650-2
2	30	18.9	4771	2	US-09-021-287-2
3	30	18.9	4771	4	US-09-240-473-2
4	30	18.9	33529	4	US-09-144-085-3
5	29.4	18.5	3147	2	US-08-781-802-7
6	29.4	18.5	3147	4	US-08-694-078-7
7	29.4	18.5	3147	4	US-09-058-260-7
8	28.6	18.0	2861	4	US-08-482-073-10
9	28.6	18.0	3647	1	US-07-914-281-7
10	28.6	18.0	3647	1	US-08-393-246-7
11	28.6	18.0	3647	2	US-08-525-058A-7
12	28.6	18.0	3647	2	US-08-696-731-7
13	28.6	18.0	3647	4	US-09-042-531-7
14	28	17.6	1146	1	US-08-482-385A-1
15	28	17.6	2728	1	US-08-482-385A-5
16	27.8	17.5	49136	4	US-09-422-869-1
17	27.6	17.4	318	2	US-08-646-981-4
18	27.6	17.4	1257	3	US-08-640-906-1
19	27.6	17.4	1257	4	US-09-395-936-1
20	27.6	17.4	2001	4	US-09-422-869-7
21	27.4	17.2	733	3	US-09-082-092-15
22	27.4	17.2	6176	3	US-08-911-321-6
23	27.4	17.2	4411529	4	US-09-103-840A-1
24	27.2	17.1	750	3	US-09-010-809-12
25	27.2	17.1	1633	4	US-09-119-788-1
26	27.2	17.1	1864	2	US-08-673-388-9
27	27.2	17.1	1864	2	US-08-614-877-9

c	28	27.2	17.1	3300	2	US-08-928-692-29	Sequence 29, Appl
c	29	27	17.0	3647	5	PCT-US91-00899-4	Sequence 4, Appl
	30	26.6	16.7	1238	2	US-08-330-290-11	Sequence 11, Appl
	31	26.4	16.6	884	2	US-08-901-200A-11	Sequence 11, Appl
	32	26.4	16.6	884	3	US-09-219-391-11	Sequence 11, Appl
	33	26.4	16.6	1723	1	US-07-841-646-28	Sequence 28, Appl
	34	26.4	16.6	1723	1	US-07-901-703-10	Sequence 10, Appl
	35	26.4	16.6	1723	1	US-08-147-023-28	Sequence 28, Appl
	36	26.4	16.6	1723	1	US-08-206-864-3	Sequence 3, Appl
	37	26.4	16.6	1723	1	US-08-278-729A-20	Sequence 20, Appl
	38	26.4	16.6	1723	1	US-08-480-528A-7	Sequence 7, Appl
	39	26.4	16.6	1723	1	US-08-479-666-7	Sequence 7, Appl
	40	26.4	16.6	1723	1	US-08-155-343A-20	Sequence 20, Appl
	41	26.4	16.6	1723	1	US-08-406-672-20	Sequence 20, Appl
	42	26.4	16.6	1723	1	US-08-643-563A-20	Sequence 20, Appl
	43	26.4	16.6	1723	1	US-08-447-570-28	Sequence 28, Appl
	44	26.4	16.6	1723	1	US-08-643-763A-20	Sequence 20, Appl
	45	26.4	16.6	1723	1	US-08-462-623-20	Sequence 20, Appl

## ALIGNMENTS

RESULT 1

US-08-866-650-2

; Sequence 2, Application US/08866650

; Patent No. 5939321

; GENERAL INFORMATION:

; APPLICANT: Greenspan, Daniel S

; APPLICANT: Takahara, Kazuhiko

; APPLICANT: Hoffman, Guy G

; TITLE OF INVENTION: Mammalian Tolloid-Like Protein

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Quarles & Brady

; STREET: 1 South Pinckney Street

; CITY: Madison

; STATE: WI

; COUNTRY: US

; ZIP: 53703

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION NUMBER: US/08/866,650

; FILING DATE:

; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:

; NAME: Berson, Bennett J

; REGISTRATION NUMBER: 37094

; REFERENCE/DOCKET NUMBER: 960296.93839

; TELEPHONE: 608-251-5000

; TELEFAX: 608-251-9166

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 4771 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 611..3652

; OTHER INFORMATION: /product= "murine mtll protein"

US-08-866-650-2

Query Match 18.9%; Score 30; DB 2; Length 4771;  
Best Local Similarity 52.4%; Pred. No. 3.2;  
Matches 66; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

### RESULT

US-09-240-473-2  
: Sequence ? Application US/09240473

APPLICANT: HOLLIMAN, Guy S  
TITLE OF INVENTION: Mammalian Tolloid-Like Protein

NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Quarries & Bricky  
STREET: 1 South Pinckney Street

CITY: MADISON  
STATE: WI

COUNTRY: US  
PID: E2702

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

```

; SOFTWARE: PatentIn Release #1.0, Version #1.30

```

; APPLICATION NUMBER: US/09/240,473

CLASSIFICATION:

NAME: Berson, Bennett J

REFERENCE/DOCKET NUMBER: 960296.93839

TELEPHONE: 608-251-5000

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:  
LENGTH: 4771 base pairs

TYPE: nucleic acid  
STRANDEDNESS: double

TOPOLOGY: Linear  
MOLECULE TYPE: DNA (genomic)

```
;
;
FEATURE:
NAME /KEY. CDS
```

```

; LOCATION: 611..3652
; OTHER INFORMATION: /product= "murine mtl protaf

```

US-09-240-473-2

1. The first step is to identify the problem or question that needs to be answered. This involves understanding the context and the specific requirements of the task.

Best Local Similarity 52.4%; Pred. No. 3.2;

201 TCGGCCGGGGTCTGACAGCGCGGGGCGGGCGGGAGCCG

Ov 123 acetata 128

DB 267 CCTCAG 266

## RESULT 4







```

; FILING DATE: 18-DEC-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/359516
; FILING DATE: 01-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/354151
; FILING DATE: 28-APR-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: B124CIP4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 596-9000
; TELEFAX: (212) 596-9090
; TELEX: 14-8367
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2861 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-482-073-10

```

```

Query Match      18.0%; Score 28.6; DB 4; Length 2861;
Best Local Similarity 57.1%; Pred. No. 7.5;
Matches 52; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

```

```

QY 5 tgaccagccacggcgctccctgtgtcagctgtgcccgtgtcccttcagggctcccgagc 64
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 340 TGAGCTGGCCAGGACGCCAACCCGGGACCGCCGCTTCTTCCACTGCGCCCGGG 281

QY 65 cacacgctggcgctgtgctgagggaaacat 95
    || || || || || || || || || || || || || || || || || || || ||
DB 280 CCCAGCCGCGCCGACGACCGCCCGGGAGCCT 250

```

```

RESULT 9
US-07-914-281-7/c
; Sequence 7, Application US/07914281
; Patent No. 5324663
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
; TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESS: P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 19920720
; APPLICATION NUMBER: US/07/914,281
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavalleye, Jean-Paul M. P.
; REGISTRATION NUMBER: 31,451
; REFERENCE/DOCKET NUMBER: 2363-060-55
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAT UR

```

```

; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3647 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; ANTI-SENSE: NO
US-07-914-281-7

```

```

Query Match      18.0%; Score 28.6; DB 1; Length 3647;
Best Local Similarity 57.1%; Pred. No. 8;
Matches 52; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

```

```

QY 5 tgaccagccacggcgctccctgtgtcagctgtgcccgtgtcccttcagggctcccgagc 64
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1733 TGAGCTGGCCAGGACGCCAACCCGGGACCGCCGCTTCTTCCACTGCGCCCGGG 1674

QY 65 cacacgctggcgctgtgctgagggaaacat 95
    || || || || || || || || || || || || || || || || || || || ||
DB 1673 CCCAGCCGCGCCGACGACCGCCCGGGAGCCT 1643

```

```

RESULT 10
US-08-393-246-7/c
; Sequence 7, Application US/08393246
; Patent No. 5595900
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION
; TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESS: P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/393,246
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/220,433
; FILING DATE: 30-MAR-1994
; APPLICATION NUMBER: US 07/914,281
; FILING DATE: 20-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Lavalleye, Jean-Paul M. P.
; REGISTRATION NUMBER: 31,451
; REFERENCE/DOCKET NUMBER: 2363-060-55
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAT UR

```

```

; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3647 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; ANTI-SENSE: NO

```



GENERAL INFORMATION:  
APPLICANT: LOWE, JOHN B.  
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS  
OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,  
TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION  
TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: P.C.  
STREET: 1755 Jefferson Davis Highway, Fourth Floor  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/042,531  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/393,246  
FILING DATE:  
APPLICATION NUMBER: US 08/220,433  
FILING DATE: 30-MAR-1994  
APPLICATION NUMBER: US 07/914,281  
FILING DATE: 20-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Lavalleye, Jean-Paul M. P.  
REGISTRATION NUMBER: 31,451  
REFERENCE/DOCKET NUMBER: 2363-060-55  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)521-4500  
TELEFAX: (703)486-2347  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3647 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA (genomic)  
ANTI-SENSE: NO  
US-09-042-531-7

Query Match 18.0%; Score 28.6; DB 4; Length 3647;  
Best Local Similarity 57.1%; Pred. No. 8;  
Matches 52; Conservative 0; Mismatches 39; Indels 0; Gaps 0;  
Qy 5 tgaccagccagggcgctcctgtgtcagctgtgcccgttccttcagggctccagcagc 64  
Db 1733 TGAGCTGGCCAGGACGGCCACCGCCGCGCTCCCTTTCTTCCACTGCGCCGGG 1674  
Qy 65 cacacgtggcgctgctggctgagggacat 95  
Db 1673 CCCAGCGCCCGCCAGGCGCCCGGAGCCT 1643

RESULT 14  
US-08-482-385A-1/c  
Sequence 1, Application US/08482385A  
Patent No. 5728561  
GENERAL INFORMATION:  
APPLICANT: DENOYA,, CLAUDIO D.  
TITLE OF INVENTION: GENES ENCODING BRANCHED CHAIN ALPHA  
TITLE OF INVENTION: KETOACID DEHYDROGENASE FROM STREPTOMYCES SPECIES  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:

ADDRESSEE: PETER C. RICHARDSON  
STREET: 235 EAST 42ND STREET, 20TH FLOOR  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: U.S.A.  
ZIP: 10017-5755  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,385A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: SHEYKA,, ROBERT F.  
REGISTRATION NUMBER: 31,304  
REFERENCE/DOCKET NUMBER: PC8346C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-573-1189  
TELEFAX: 212-573-1939  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1146 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-482-385A-1

Query Match 17.6%; Score 28; DB 1; Length 1146;  
Best Local Similarity 55.0%; Pred. No. 8.7;  
Matches 55; Conservative 0; Mismatches 45; Indels 0; Gaps 0;  
Qy 15 cggcgctccctgctgcagctgtgccttcagggtcccgagccagcacgctgg 74  
Db 224 CGGCCCTGCTTGGTGAGAGCCGTGGCTGGTGTGTAGCGGGCGCGCCAGCTCC 165  
Qy 75 gcgtgctggctgagggaacatggcattgttgccctcagctg 114  
Db 164 GCGTACAGCGCGGCAGCAGTAGCGGGTGGCTCCGCGG 125

RESULT 15  
US-08-482-385A-5/c  
Sequence 5, Application US/08482385A  
Patent No. 5728561  
GENERAL INFORMATION:  
APPLICANT: DENOYA,, CLAUDIO D.  
TITLE OF INVENTION: GENES ENCODING BRANCHED CHAIN ALPHA  
TITLE OF INVENTION: KETOACID DEHYDROGENASE FROM STREPTOMYCES SPECIES  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PETER C. RICHARDSON  
STREET: 235 EAST 42ND STREET, 20TH FLOOR  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: U.S.A.  
ZIP: 10017-5755  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/482,385A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: SHEYKA,, ROBERT F.  
REGISTRATION NUMBER: 31,304

```

; REFERENCE/DOCKET NUMBER: PC8346C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-573-1189
; TELEFAX: 212-573-1939
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2728 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-482-385A-5

```

```

Query Match      17.6%; Score 28; DB 1; Length 2728;
Best Local Similarity 55.0%; Pred. NO. 11;
Matches 55; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

Qy 15 cggggtccctgctgcagctgcccgcctccaggctcccgagccacacgctgg 74
Db 626 CGGCCCTGCTTGTGAGAGCCGTGGCTGCTGTTGTAGCGCGCGCGCCACCACTCC 567

Qy 75 gcgtgctggctgaggggaacatggcatgttgccctcagctg 114
Db 566 GCGTACAGCGCGGCACGACGTAGCGGGTGGCTCGCGCG 527

```

Search completed: September 20, 2002, 03:22:05  
Job time: 14214 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 01:23:18 ; Search time 3895 Seconds  
(without alignments)  
550.967 Million cell updates/sec

Title: US-09-846-456-5  
Perfect score: 159  
Sequence: 1 ttaatgaccgcacggcg.....ctttcagaagaagacaaca 159

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	140.6	88.4	736	9	AU135588
2	72.8	45.8	619	9	BB657864
3	70.8	44.5	292	10	Z44377
4	70.8	44.5	535	10	BG384217
5	37	23.3	440	9	AA914462
6	36	22.6	533	9	AV845237
7	35.6	22.4	514	12	BH087817
8	35.4	22.3	482	12	AQ848626
9	35	22.0	571	9	AV862022
10	34.8	21.9	301	9	AW751638
11	34.8	21.9	365	9	BB843149
12	34.8	21.9	516	10	BF484412
13	34.8	21.9	530	10	BE471178
14	34.6	21.8	344	10	BM149133
15	34.6	21.8	1574	10	BG325753
16	34.4	21.6	571	9	AV864011
17	34.4	21.6	579	9	AV892280

18	34	21.4	598	12	BH087727
19	33.8	21.3	668	10	BG398045
20	33	20.8	335	9	AW313960
21	33	20.8	477	12	AQ885360
22	33	20.8	575	10	BI541694
23	33	20.8	791	10	BG675383
24	33	20.8	1901	10	BF128237
25	32.8	20.6	543	10	BG747266
26	32.8	20.6	997	12	CNS006DN
27	32.6	20.5	653	9	BB132336
28	32.6	20.5	870	9	AL530607
29	32.4	20.4	658	9	BB866093
30	32.4	20.4	667	9	BB653232
31	32.2	20.3	684	10	BG700305
32	32.2	20.3	715	10	BG717361
33	32.2	20.3	971	10	BI915421
34	32	20.1	277	9	BB265873
35	32	20.1	650	9	AL636346
36	32	20.1	771	10	BI770317
37	32	20.1	777	10	BG830641
38	32	20.1	884	10	BI668321
39	32	20.1	935	9	AL539815
40	31.8	20.0	548	12	AQ452008
41	31.8	20.0	615	10	BE981503
42	31.8	20.0	682	9	BB644307
43	31.8	20.0	724	10	BE981359
44	31.8	20.0	801	12	AQ745776
45	31.6	19.9	472	9	AA726802

## ALIGNMENTS

RESULT 1  
AU135588  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
FEATURES  
BASE COUNT  
ORIGIN

AU135588  
AU135588  
sequence.  
AU135588  
AU135588.1  
EST.  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 736)  
Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and Isogai, F.  
HRI human cDNA project  
Unpublished (2000)  
Contact: Takao Isogai  
Genomics Laboratory  
Helix Research Institute  
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
Tel: 81-438-52-3951  
Fax: 81-438-52-3952  
Email: genomics@hri.co.jp  
HRI human cDNA project; 5' - & 3' - end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.  
Location/Qualifiers  
1. .736  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="PLACE1002437"  
/clone\_lib="PLACE1"  
/tissue\_type="placenta"  
/notes="Vector: pME18SFL3"  
163 a 199 c 199 g 170 t  
5 others

736 bp mRNA linear EST 24-OCT-2000  
PLACE1002437 5', mRNA

GI:10996127

FEATURES	Source
e mouse	tissues.
	Location/Qualifiers
	1. 619
	/organism="Mus musculus"
	/db_xref="taxon:10090"
	/clone="p230019p04"
	/clone_lib="RIKEN full-length enriched, 12 days embryo
	evball"

/tissue\_type="eyeball"  
 /dev\_stage="12 days embryo"  
 /lab\_host="DH10B"  
 /note="Site\_1: Salt; Site\_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5',  
 GAGAGAGACGGCGCGCAACTGTTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the

BASE COUNT	126 a	185 c	173 g	134 t	1 others
ORIGIN					
Query Match	45.8 %; Score 72.8; DB 9; Length 619;				
Best Local Similarity	77.4 %; Score 125.5; DB 9; Length 619;				

RESULT	3	244377	292 bp	linear	EST 14-NOV-1994
OCUS		HSC1ZB081	normalized infant brain	cdna Homo sapiens	cdna clone
DEFINITION		c-12b08,	mRNA sequence.		
ACCESSION		244377			
VERSION		244377.1	GI:573506		
KEYWORDS		EST.			
ORIGIN		human.			
ORGANISM		Homo sapiens			
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS		1. (bases 1 to 292)			
		Auffray,C., Behar,G., Bois,F., Bouchier,C., da Silva,C., Devignes ,M.D., Duprat,S., Houlgatte,R., Jumeau,M.N., Lamy,B., Lorenzo,F., Mitchell,H., Mariage-Samson,R., Pietu,G., Poulliot,Y., Sebastiani-Kabaktchis,C. and Tessier,A.			
TITLE		IMAGE: molecular integration of the analysis of the human genome and its expression			
JOURNAL		C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)			
MEDLINE		95277534			
COMMENT		Contact: Genethon Genexpress-Genethon Genethon Centre de recherche sur le Genome Humain 1,rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE Tel: 33169472800			





Db 66 ATTGTGAAGCAGCATTCGATCACAAAAAGCTCATA 31

ACCESSION  
VERSION

Leishmania major genomic  
AQ848626  
AQ848626.1 GI:6053374

```

KEYWORDS  GSS.
SOURCE      Leishmania major.
ORGANISM    Leishmania major
            Leishmania.
REFERENCE   1 (bases 1 to 482)
AUTHORS     Akopyants, N.S., Clifton, S.W., Martin, J., Pape, D., Wylie, T., Li, L.,
            Kissinger, J., Roos, D.S., Marra, M., Hillier, L., Chinwalla, A., Ronko
            Blistain, A., Schmitt, A., Person, B., Theising, B., Ritter, E., Ronko
            J., Bennett, J., Cole, R., Underwood, K., Cardenas, M., Gibbons, M.,
            Harvey, N., McCann, R., Tsagareishvili, R., Williams, T., Jackson, Y.,
            Bowers, Y., Swaller, T., Waterston, R., Wilson, R. and Beverley, S.M.
TITLE       A survey of the Leishmania major Friedlin strain V1 genome by
            shotgun sequencing: a resource for DNA microarrays and expression
            profiling
JOURNAL     Mol. Biochem. Parasitol. 113 (2), 337-340 (2001)
MEDLINE     21192569
COMMENT     Other GSSs: lm12d06.y1
            Contact: Akopyants, NS / Beverley, SM
            WashU Leishmania Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: estewatson.wustl.edu
            Library construction: Natalia S. Akopyants, Ph.D.
            DNA sequencing by: Washington University Genome Sequencing Center
            If using this information please cite:
            N.S. Akopyants and S.M. Beverley 'A survey of the Leishmania major
            Friedlin strain V1 genome by shotgun sequencing' and the Washington
            University Genome Sequencing Center For information on obtaining
            clone material please contact: Natalia S. Akopyants Ph.D.
            (natalia@borcim.wustl.edu) and/or Stephen M. Beverley Ph.D.
            (beverley@borcim.wustl.edu)
            Seq primer: -40UP from Gbco
            Class: shotgun
            High quality sequence stop: 469.
FEATURES    Location/Qualifiers
            source
            1..482
            /organism="Leishmania major"
            /strain="Friedlin strain V1"
            /db_xref="taxon:5664"
            /clone="LMAJFV1_lm12d06"
            /lab_host="Leishmania major FV1 random genomic library"
            /note="Vector: pZero-2 (Invitrogen)"
            Genomic DNA was isolated from stationary phase cells. For
            this library, DNA was sheared to give a tight size
            distribution of 1-1.5kb fragments, blunt-ended with T4 DNA
            phosphatase, dephosphorylated with Shrimp Alkaline
            Phosphatase and ligated into pZero-2 vector's EcoRV site."
BASE COUNT  87 a 166 c 145 g 83 t
ORIGIN

Query Match      22.3%; Score 35.4; DB 12; Length 482;
Best Local Similarity 57.8%; Pred. No. 11;
Matches 63; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Qy 10 agccacggggtccctgctgcagctgtggccgtgcttcacaggggtccgcagccacac 69
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 231 AGGCTGTGCCGTCAGCTTGCACACTTGAAGCAGCAGCAGCTCAGCGCGCGGTGCTGGG 290

Qy 70 gctggcgctgctgctgaggaaacatggcattgttgctcagctgaggt 118
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 291 GCCCGCGGAGTACTTCATGCGCACTTCCCGTGTGACCTCAGATGCTGT 339

RESULT 9
AV862022/c
LOCUS      571 bp mRNA linear EST 08-NOV-2001
DEFINITION Nori Satoh unpublished cDNA library, egg Ciona
            intestinalis cDNA clone rcieg31o05 3', mRNA sequence.

GSS.
Leishmania major.
Leishmania major
Leishmania.
1 (bases 1 to 482)
Akopyants, N.S., Clifton, S.W., Martin, J., Pape, D., Wylie, T., Li, L.,
Kissinger, J., Roos, D.S., Marra, M., Hillier, L., Chinwalla, A., Ronko
Blistain, A., Schmitt, A., Person, B., Theising, B., Ritter, E., Ronko
J., Bennett, J., Cole, R., Underwood, K., Cardenas, M., Gibbons, M.,
Harvey, N., McCann, R., Tsagareishvili, R., Williams, T., Jackson, Y.,
Bowers, Y., Swaller, T., Waterston, R., Wilson, R. and Beverley, S.M.
A survey of the Leishmania major Friedlin strain V1 genome by
shotgun sequencing: a resource for DNA microarrays and expression
profiling
Mol. Biochem. Parasitol. 113 (2), 337-340 (2001)
21192569
Other GSSs: lm12d06.y1
Contact: Akopyants, NS / Beverley, SM
WashU Leishmania Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
Library construction: Natalia S. Akopyants, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
If using this information please cite:
N.S. Akopyants and S.M. Beverley 'A survey of the Leishmania major
Friedlin strain V1 genome by shotgun sequencing' and the Washington
University Genome Sequencing Center For information on obtaining
clone material please contact: Natalia S. Akopyants Ph.D.
(natalia@borcim.wustl.edu) and/or Stephen M. Beverley Ph.D.
(beverley@borcim.wustl.edu)
Seq primer: -40UP from Gbco
Class: shotgun
High quality sequence stop: 469.

Accession      AV862022
Version        AV862022.1
Keywords       EST.
Source         Ciona intestinalis.
Organism       Ciona intestinalis
              Ciona intestinalis
              Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
              Phlebobranchia; Clonidae; Ciona.
              1 (bases 1 to 571)
              Satoh, N., Satoh, Y., Kohara, Y. and Shin-i, T.
              Expressed genes in Ciona intestinalis
              Unpublished (2000)
              Contact: Nori Satoh
              Department of Zoology
              Kyoto University
              Sakyo-Ku, Kyoto, Kyoto 606-8502, Japan
              Tel: 81-75-753-4081
              Fax: 81-75-705-1113
              Email: satoheascidian.zool.kyoto-u.ac.jp.
              Location/Qualifiers
              1..571
              /organism="Ciona intestinalis"
              /db_xref="taxon:7719"
              /clone="rcieg31o05"
              /clone_lib="Nori Satoh unpublished cDNA library, egg"
              /tissue_type="whole animal"
              /dev_stage="egg"
BASE COUNT  154 a 116 c 99 g 201 t 1 others
ORIGIN

Query Match      22.0%; Score 35; DB 9; Length 571;
Best Local Similarity 51.3%; Pred. No. 14;
Matches 80; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

Qy 4 atgaccagccacggggtccctgctgcagctgtgcccgtgcttcacaggggtccgcag 63
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 211 ATGAGAGGAGCAAGATCTGCATGTTCTGTACACTGAACCTGCTTTTCATTGGATCAAAAT 152

Qy 64 ccacagctggggtgctgctgagggagacatgcatgtgtgctcagctgaggtgtgtg 123
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 151 GCCAACGCTGTGCCAGCAGTAAGAAGAAATATGGAATGCCAGTCAATTTGCAGAAATGTA 92

Qy 124 ctgtggaagaacctcacttccagaaagaagacaaaca 159
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 91 ATTTGAACGACGATTTNATCACAAAAAAGCTCAT 56

RESULT 10
AW751638/c
LOCUS      301 bp mRNA linear EST 28-APR-2000
DEFINITION PC5-CT0071-041099-002-g07 CT0071 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW751638
VERSION     AW751638.1
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 301)
            Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
            Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
            Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
            Brustein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
            M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
            Simpson, A.J.
            Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
            Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
            20202663
            Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

```



US Department of Agriculture, Agriculture Research Service, Pacific  
West Area, Western Regional Research Center  
800 Buchanan Street, Albany, CA 94710, USA  
Tel: 5105595773  
Fax: 5105595818  
Email: oanderns@pw.usda.gov

Sequence have been trimmed to remove vector sequence and low  
quality sequence with phred score less than 20  
Seq primer: Stratagene SK primer.

# FEATURES

source

Location/Qualifiers

1. .516

/organism="Triticum aestivum"

/cultivar="Chinese Spring"

/db\_xref="taxon:4565"

/clone="WHE2323\_B07\_D13"

/clone\_lib="Wheat pre-anthesis spike cDNA library"

/tissue\_type="Spike before anthesis"

/dev\_stage="Adult plant"

/lab\_host="E. coli SOLR"

/note="Vector: Lambda Uni-ZAP XR, excised phagemid;

Site\_1: EcoRI; Site\_2: XhoI; plants were grown in the

greenhouse. Whole spike with awns trimmed, white, green

and yellow anther were collected and total RNA, and

poly(A) RNA were prepared, a cDNA library was made, and

the cDNA clones were in vivo excised to give pBluescript

phagemids in the TJ Close lab (Choi, Close, Fenton) at

the University of California, Riverside. Plasmid DNA

preparations and DNA sequencing were performed in the OD

Anderson lab (all other authors)."

BASE COUNT 106 a 177 c 136 g 97 t

ORIGIN

Query Match 21.9%; Score 34.8; DB 10; Length 516;  
Best Local Similarity 57.3%; Pred. No. 16;  
Matches 63; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 19 cgtccctgtcagctctgcccgtctccagggtcccgagccacacgtggcggt 78

Db 276 CGCCGCTGCCGACATCTCCGACCCATGACGAGCAGCTCAGTCAAGACAGATGTCGGCGAG 335

QY 79 gctggtgagggaacatggcatgttgctcagctcagctgaggtgtgtgtg 128

Db 336 GGGAGAAAAGGAAGATGGGATCGCGCATCATCTCTGGACGCCGCTGTG 385

## RESULT 13

BE471178 530 bp mRNA linear EST 28-JUL-2000

LOCUS WHE0285\_G07\_N13ZS Wheat drought-stressed seedling cDNA library

DEFINITION Triticum aestivum cDNA clone WHE0285\_G07\_N13, mRNA sequence.

ACCESSION BE471178

VERSION BE471178.1 GI:9561669

KEYWORDS EST.

SOURCE bread wheat.

ORGANISM Triticum aestivum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Authors Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae

1 (bases 1 to 530)

Anderson, O.D., Chao, S., Choi, D.W., Close, T.J., Fenton, R.D., Han

, P.S., Hsia, C.C., Kang, Y., Lazo, G.R., Miller, R., Rausch, C.J.,

Seaton, C.L. and Tong, J.C.

The structure and function of the expressed portion of the wheat

genomes - Drought-stressed seedling cDNA library

Unpublished (2000)

Contact: Olin Anderson

US Department of Agriculture, Agriculture Research Service, Pacific

West Area, Western Regional Research Center

800 Buchanan Street, Albany, CA 94710, USA

Tel: 5105595773

Fax: 5105595818

Email: oanderns@pw.usda.gov

Sequence have been trimmed to remove vector sequence and low  
quality sequence with phred score less than 20  
Seq primer: Stratagene SK primer.

# FEATURES

source

Location/Qualifiers

1. .530

/organism="Triticum aestivum"

/cultivar="Chinese Spring"

/db\_xref="taxon:4565"

/clone="WHE0285\_G07\_N13"

/clone\_lib="Wheat drought-stressed seedling cDNA library"

/tissue\_type="Seedling without endosperm"

/dev\_stage="Five day old seedling"

/lab\_host="E. coli SOLR"

/note="Vector: Lambda Uni-ZAP XR, excised phagemid;

Site\_1: EcoRI; Site\_2: XhoI; Seeds were surface-sterilized

, germinated and grown aseptically in the dark at room

temperature on filter paper with water, nystatin and

cefotaxime in covered crystallization dishes. Five-day old

seedlings were incubated for one day at 90% RH. After

removing endosperm, seedlings were transferred to

desiccator jar containing saturated MgSO4 at room

temperature for 24 hr. The tissue, total RNA, and poly(A)

RNA were prepared, a cDNA library was made, and the cDNA

clones were in vivo excised to give pBluescript phagemids

in the TJ Close lab (Choi, Close, Fenton) at the

University of California, Riverside. Plasmid DNA

preparations and DNA sequencing were performed in the OD

Anderson lab (all other authors)."

BASE COUNT 113 a 178 c 143 g 96 t

ORIGIN

Query Match 21.9%; Score 34.8; DB 10; Length 530;  
Best Local Similarity 57.3%; Pred. No. 16;  
Matches 63; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 19 cgtccctgtcagctctgcccgtctccagggtcccgagccacacgtggcggt 78

Db 275 CGCCGCTGCCGACATCTCCGACCCATGACGAGCAGCTCAGTCAAGACAGATGTCGGCGAG 334

QY 79 gctggtgagggaacatggcatgttgctcagctcagctgaggtgtgtgtg 128

Db 335 GGGAGAAAAGGAAGATGGGATCGCGCATCATCTCTGGACGCCGCTGTG 384

## RESULT 14

BM149133

LOCUS

DEFINITION

BM149133

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BM149133 344 bp mRNA linear EST 30-NOV-2001

LOCUS TCAAP2E6391 Pediatric acute myelogenous leukemia cell (FAB M1)

DEFINITION Baylor-HGSC project-TCAA Homo sapiens cDNA clone TCAAP6391, mRNA

sequence.

ACCESSION BM149133

VERSION BM149133.1 GI:17170474

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Authors Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

1 (bases 1 to 344)

Wei, Y., Tsang, Y.T.M., Mei, G., Ku, J.M., Ali-Osman, F.R. Jr.,

Gunarathne, P.H., Muzny, D., Bouck, J., Gibbs, R.A. and Margolin, J.F.

Pediatric Leukemia cDNA Sequencing Project (2001)

Unpublished (2001)

Contact: Dr. Judith F. Margolin

Texas Children's Cancer Center and Human Genome Sequencing Center

at Baylor College of Medicine

1102 Bates, MC3-3320 Houston, TX 77030, USA

Tel: 832-824-4536

Fax: 832-825-4038

Email: clones@txccc.org

Seq primer: M13 primer.

FEATURES

source

Location/Qualifiers

1. .344

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="TCRAP6391"
/clone_lib="pediatric acute myelogenous leukemia cell (FAB M1) Baylor-HGSC project-PCAA"
/sex="male"
/tissue_type="leukopheresis"
/cell_type="myeloid cell"
/dev_stage="pediatric 6 years"
/lab_host="DH10B"
/Note="Vector: lambda pSB; Site_1: BamHI; Site_2: EcoRI; First strand cDNA was primed with an anchored XhoI-oligo(dT) primer [5'GGAGGACTGAGCGCGCAGGAGAG(T)VN 3'; V=A,C,G; N=A,C,G,T] and then dG tailed. Second strand was primed with a BamHI-dC primer [5'AGAGGCTGGATCGCGCGCCCAATATAAT(C) 3']. Double-stranded cDNA was then digested with BamHI and XhoI and directionally cloned into the BamHI and SalI sites of lambda pSB vector. Library was constructed by Wei Yu at RIKEN of Japan (Carninci P, Westover A, Nishiyama Y, Ohsumi T, Itoh M, Nagasaki S, Sasakura, Okazaki Y, Muramatsu M, Schneider C, Hayashizaki Y, High efficiency selection of full-length cDNA by improved biotinylated cap trapper., DNA Res 4: 1, 61-6, Feb 28, 1997)"
BASE COUNT      57 a 108 c 106 g 73 t
ORIGIN

```

```

Query Match      21.8%; Score 34.6; DB 10; Length 344;
Best Local Similarity 70.8%; Pred. No. 16;
Matches 46; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 91 aacatgcatgttgccctcagctgaggtgctgctggaagaaccttcacgaaga 150
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 232 ACATGCGCTTCGGACACAGCTGATGCTGCTCTGGAAGAATTCATGATGCCGG 291
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 151 agaca 155
    |||||
DB 292 AGACA 296

```

## RESULT 15

```

BG325753
LOCUS      60242464F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4562559 5',
DEFINITION mRNA sequence.
ACCESSION BG325753
VERSION    BG325753.1 GI:13132177
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 1574)
            NIH-MGC http://mgi.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTF
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone Distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LICM1274 row: p column: 16
            High quality sequence start: 5
            High quality sequence stop: 178.
FEATURES
    source      1..1574
                /organism="Homo sapiens"

```

```

/db_xref="taxon:9606"
/clone="IMAGE:4562559"
/clone_lib="NIH_MGC_14"
/tissue_type="renal cell adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: kidney; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
```

```

BASE COUNT      305 a 588 c 432 g 249 t
ORIGIN

Query Match      21.8%; Score 34.6; DB 10; Length 1574;
Best Local Similarity 55.4%; Pred. No. 25;
Matches 67; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 2 taatgaccagccacggcgctccctgctcagctcgtgcccgtcctccagggctccgcg 61
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1216 TCACGAGCCCCCTTGTTCCTGCTGCGCTCTTTCCGCTGCTGCTGTCGCGCGG 1275
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 62 agccacacgctggcgctgctgctgaggggaacatggcatgttgccctcagctgaggttgc 121
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1276 AGCCAAAGCGCACGGAGCGGCTCTGCGAGTAGGAGAGCTTGCGCCCGCGGGTGGAGC 1335
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 122 t 122
DB 1336 T 1336

```

Search completed: September 20, 2002, 01:23:20  
Job time: 7380 sec







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 01:23:26 : Search time 5250.46 Seconds  
(without alignments)  
12877.660 Million cell updates/sec

Title: US-09-846-456-1  
Perfect score: 3231  
Sequence: 1 acaggcgtgtggcagtg.....gcccacatccaccactt 3231

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0  
Searched: 1797656 seqs, 10463268293 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_bhg.\*
- 3: gb\_in.\*
- 4: gb\_ov.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_ov.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htgo\_inv.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
-----					

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

## ALIGNMENTS

RESULT	1	AX351029	3231 bp	DNA	linear	PAT 06-FEB-2002
LOCUS	AX351029	Sequence 1 from Patent WO0183746.				
DEFINITION	AX351029	Sequence 1 from Patent WO0183746.				
ACCESSION	AX351029	GI:18616385				
VERSION	AX351029.1					
KEYWORDS						
SOURCE		human.				
ORGANISM		Homo sapiens				
REFERENCE		Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
AUTHORS		1 (sites)				
TITLE		Rosier-Montus, M.F., Prades, C., Lemoine, C., Naudin, L., Deneffe, P., Brewer, B., Duverger, N., Remaley, A. and Santamarina-Fojo, S.				
JOURNAL		Regulatory nucleic acid sequences of the abcl gene				
FEATURES		Patent: WO 0183746-A 1 08-NOV-2001;				
		Aventis Pharma S.A. (FR)				
		Location/Qualifiers				
		1..3231				
		/organism="Homo sapiens"				
		/db_xref="taxon:9606"				
BASE COUNT		809 a	773 c	876 g	773 t	
ORIGIN						

Query Match		100.0%;	Score 3231;	DB 6;	Length 3231;		
Best Local Similarity		100.0%;	Pred. No. 0;				
Matches 3231;		Conservative	0;	Mismatches	0;	Indels	0; Gaps
		0;					
Qy	1	acagggcattgagcagctgtaattcagttactcgggaggttggaattgcaatga	60				
Db	1	ACAGGGCATGGTGCAGGTGCTTAATCTCAGTTACTCGGGAGGTGGAGTTGCAATGA	60				
Qy	61	gccagatcgaccattgacactccagctgggcaaaaaagggtgaactccatccaatt	120				
Db	61	GCCAGATCGACCATTTGCACTCCAGCTCGGCAACAAAAGGTGAACACTCCATCTCAATT	120				
Qy	121	aaaaaaaagaattgatttgggtcgacttcaaatagtagagaaagaagagagagg	180				
Db	121	AAAAAAAAGATGATTTTGGTGGTGCATTTCAATAGTAGAGGAAGAGGAGAGG	180				
Qy	181	agatggaggtcgaggagatctataactctctataatcatgctaggaagataaacacct	240				
Db	181	AGATGGAGGGTCAGGGAGATCTAATTACTCTCTAAAATCATGCTAGGAAGATAACACCT	240				
Qy	241	ttaataaacactctgcttttataacatcatctcgccaggagctcaagggtttcaaca	300				
Db	241	TTTAATAACACTCTCTGCTTTTATAACATCATTTCTCCAAAGAGCTCAAGGTTTCAAGA	300				
Qy	301	aagttcacttccagaaaccccttgaggagacagaaatatacatctctctccatttta	360				
Db	301	AAGTTCACTTTCAGAAACCCCTTTGAGGAAGACAAATATACATCTTCTCTCCATTTTA	360				
Qy	361	aagatgaagaacaggccgggacaaatggctaatgctgtaatccccagcaccttgggagg	420				
Db	361	AAGATGAAGAAGAACAGGCGGGCAATGGCTAATGCTGTAAATCCAGCACCTTTGGGAGG	420				
Qy	421	ctaagccagagatcgcttgactccagagtttgacagcagctggataaactggcaaa	480				
Db	421	CTAGGCCAGAGGATCGCTTGAGCTCCAGAGTTTGAGACAGCCTGGATTAACATGGCAAA	480				
Qy	481	acctgtctctacaaaaataacaaaaattagatgggtggtggtgcacgtgacctgtgt	540				
Db	481	ACCTGTCTCTACAAAAAATAACAAAAATTAGATGGGTGTGGTGCAATGCACCTGTGTT	540				
Qy	541	ccagctacttgggagctaaagtgaggagatcgcttgagccagggagtgcaagtctaca	600				
Db	541	CCAGCTACTTTGGGAGGCTTAAGTGGGAGGATCGCTTGAGCCAGGAGTCAAGTCTACA	600				
Qy	601	ctgagccatgattggatcactgcactccagcctgggtggtgacagagcaaacctgtctca	660				
Db	601	CTGAGCCATGATTGGATCAGTCTGACCTCCAGCTGGGTGAGTACAGAGCAAGCCTGTCTCA	660				
Qy	661	aaaaaagaatgaagagaaagaagaagagagagagagagagagagagagagagag	720				
Db	661	AAAAAAGAAATGAAGAGAAAGAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG	720				
Qy	721	gagggggag	780				
Db	721	GAGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	780				
Qy	781	gatgaacag	840				
Db	781	GATGAACAG	840				
Qy	841	ccccaaacccaaattattgacaaaggttattcttgactgaggaaggggctccgctct	900				
Db	841	CCCCAAACCCAAATTTATTGACCAAGGTTATTCTTTGACTGAGGCAAGGGGCTCCGCTCT	900				
Qy	901	ctgggcttgggctttgaaagctcatctcgcccttctgagatccatccattctttt	960				
Db	901	CTGGGCTTGGGCTTTAGAAAGCTCATCTCGCCCTTTCTGAGATCCATCCCTTCTTT	960				
Qy	961	ttaattttttgacagggagttgctctgtcactcaggttgagtgagtgagtgagtc	1020				
Db	961	TTAATTTTCTTGACACGGAGCTTGCTCTGTCTCAGCTCAGGCTGGAGTGCAGTGCATGATC	1020				

Qy	1021	tcgactcactgtaaacctctgctccgggttcaagcagattctctcctcagctccctcga	1080
Db	1021	TGACTCAGTAAACCTCTGCCTCCGGGTTCAAGCGATTCTCTCCCTCAGCTCCCTGA	1080
Qy	1081	gatacagggcccccaccacacatctggctaatttttatttttagtaaaagctgggtt	1140
Db	1081	GATAACAGCGCCCCACCACACATCTGGCTAATTTTGTATTTTATTTAGTAAAGACTGGGTT	1140
Qy	1141	tcataatgttgcccaagggttgggttcgaactcctgacctgaggtgagctgccacattggc	1200
Db	1141	TCATCATGTTGGCCAGGTTGGTTTGGAACTCTGACCTCAGGTGAGCTGCCACCTTGGC	1200
Qy	1201	ctcccaaatgctggattacagcagcatagccactgcgccagctcaatccatttt	1260
Db	1201	CTCCCAAGTGTGGGATTACAGGCATGAGGCACCTGCGCCACGCTCAGATCATCCCCTTT	1260
Qy	1261	ctaaagggcaaacagtcctatgggtgcaaggggccaatgcccacccagagttatgagtacctgg	1320
Db	1261	CTAAGGGCAAAACAGTCCATGTTGTCGAAAGGGGCCATGCCACCCAGAGTTATGAGTACCTGG	1320
Qy	1321	gactccgaattccttgctgggtggcctccacatgcactccagggcctgcttgggcttc	1380
Db	1321	GACTCCGAATTTCTTGGCTGGTGGCTCCACATGCACTCCAGGGCCTGCTTGGGCTC	1380
Qy	1381	ttctatgctgctcctgagttgattgaaccactgatgtagtacctggcttgagcc	1440
Db	1381	TTCTATGGCTCTGCTCCTGAGTGTGATGAACCACTGATGAGTACCTGGGCTTGAACC	1440
Qy	1441	gtggcctggagatcctgttgactgtagcatgaggggcttctgagctgaatgctctga	1500
Db	1441	GTGGCTGGAGATCCTGTTGACTGTAGCATGGAGGGGCTTGTGCAGCTGAATGCTGCA	1500
Qy	1501	tgcaaggtgggtgagttctggaatatgtagagctggagggtgggaagaagtaggcttg	1560
Db	1501	TGCAAGGTGGTGGGAGTTCTTGAATATGATGGAGCTGGAGGTGGGAAGAAGTAGGCTTG	1560
Qy	1561	gggcagctctctcatgcccactcattctggccaaactcaggtcaaacgtgtaagagctct	1620
Db	1561	GGGCAGCTCTCTCATGCCACCTCATCTTGGCCAAACTCAGGTCAAACTGTGAAGAGTCT	1620
Qy	1621	aaatgtaatcgccttcaagggtggctcaagagatctttgtcaaggtgagagacctt	1680
Db	1621	AAATGTGAATCTGCCCTTCAAGGTGGCTTCAAGAGATATCTTTGTCAAGGTAGGAGACCTT	1680
Qy	1681	gtggcctccagctgcaactccagggcctgctgggcctcttctacagggctctgctgagt	1740
Db	1681	GTGGCTCCAGCTGCATTCAGGGGCTTGTGGGCTTCTTCTACGGGCTCTGTCTGAGT	1740
Qy	1741	cttctatgaatccttcagggcagattctatattagactcttcacagtttgacctgagttt	1800
Db	1741	CTTCTATGAATCTTCAGGGCAGATTCTATATTAGACTCTTTCACAGTTTGACCTGAGTTT	1800
Qy	1801	tggccagaataaggtgacatttagttgtgcttgatggatgacttaaatatttagaca	1860
Db	1801	TGGCCAGAATAAGGTGACATTTAGTTTGTGGCTTGATGGATGACTTTAAATATTAGACA	1860
Qy	1861	tgggtgtgagcctgcatctcactctgctcttttttttggccctccaggtgttttgggt	1920
Db	1861	TGGTGTGAGGCTTGCAATCTCTACTCTTGTCCCTTTTGTGGCCCTCCAGTGTGTTGGGT	1920
Qy	1921	agtttctcctccctacagccaaaggcaagagagaaggtggaggtctggaggtggtacat	1980
Db	1921	AGTTTCTCCTCCCTACAGCCAAAGGCAAAACAGAGAGTGGAGGCTCTGGAGTGGCTACAT	1980
Qy	1981	aattttacagagactgcaattctctggtgcaacttcaaaatgatatcaaaactaaatacaa	2040
Db	1981	AATTTTACAGACTGCAATTTCTTGGGTGCACTTTCACAAATGTATACAACTATAAATAACA	2040
Qy	2041	gtcctgtgtttttatcacagggaggtgataataataataataataataataataataata	2100
Db	2041	GTCCTGTGTTTTATCACAGGGAGGCTGATCAATATATGAAATTAAGAGGGGCTGGTC	2100







/number=9  
 125285..125424  
 /gene="ABCA1"  
 /number=10  
 125754..125870  
 /gene="ABCA1"  
 /number=11  
 130079..130276  
 /gene="ABCA1"  
 /number=12  
 131023..131228  
 /gene="ABCA1"  
 /number=13  
 131749..131925  
 /gene="ABCA1"  
 /number=14  
 133708..133930  
 /gene="ABCA1"  
 /number=15  
 135677..135898  
 /gene="ABCA1"  
 /number=16  
 136959..137163  
 /gene="ABCA1"  
 /number=17  
 138268..138381  
 /gene="ABCA1"  
 /number=18  
 140179..140350  
 /gene="ABCA1"  
 /number=19  
 141340..141471  
 /gene="ABCA1"  
 /number=20  
 142777..142919  
 /gene="ABCA1"  
 /number=21  
 143123..143260  
 /gene="ABCA1"  
 /number=22  
 143963..144183  
 /gene="ABCA1"  
 /number=23  
 145441..145513  
 /gene="ABCA1"  
 /number=24  
 146657..146859  
 /gene="ABCA1"  
 /number=25  
 148527..148575  
 /gene="ABCA1"  
 /number=26  
 148771..148884  
 /gene="ABCA1"  
 /number=27  
 150280..150428  
 /gene="ABCA1"  
 /number=28  
 152078..152202  
 /gene="ABCA1"  
 /number=29  
 153438..153536  
 /gene="ABCA1"  
 /number=30  
 156568..156757  
 /gene="ABCA1"  
 /number=31

Query Match 54.1%; Score 1749; DB 9; Length 201144;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1749; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 4 gggcattggtggcagggtgctgttaattctcaggttactcgaggagggtggaattgcaatgagcc 63

Db	31069	GGGCATGGTGGCAGGTGCCTGTAAATCTCAGTTCGGAGGTGGAGGTGCAATGAGCC	31128
Qy	64	cagatcgaccattgcactccagcctgggcaacaaggtgaaactccatctcaataaa	123
Db	31129	CAGATCGACCATTTGCATCCAGCCTGGGCAACAAAGGTGAACCTCCATCTCAATAAA	31188
Qy	124	aaaaaagaatgattttggtggtcgacttcaaatagtgaggagaagaagagagaga	183
Db	31189	AAAAAAGAATGATTTGGTGGTCGACTTCAATAGTAGGACAAGAGAGAGAGA	31248
Qy	184	tggagggtcaggagagatcctaattactctctaaatacatgctaggaagaatacacctttt	243
Db	31249	TGGAGGGTCAGGGAGATCTAAATTAATCTCTTAAATCATGCTAGGAAGATTAACACCTTTT	31308
Qy	244	ataaacactctctgcttttataacatactctccaaagagctcaaaaggttttcaacaaag	303
Db	31309	ATAAACACTCTCTGCTTTTATAACATCATCTGCCAAGGAGCTCAAAAGGTTTCAACAAAG	31368
Qy	304	ttcactttcagaaaacccctttgagaagacagaaataatactctctctocattttaag	363
Db	31369	TTCACTTTTCAGAAAACCCCTTTGAGGAAGACAGAAATATACATCTTCTCCATTTTAAAG	31428
Qy	364	atgaagaaacagccggggcaaatggttaactgcttaataatccccagcacttttggaggctg	423
Db	31429	ATGAAGAAACAGCCGGGCACAATGGCTAATGCTGTGTAATCCCCAGCACCTTTGGGAGGCTG	31488
Qy	424	agccagagatcgcttgagctccagagtttgagaccagcctgggataacaatggcaaaacc	483
Db	31489	AGCCAGAGGATCGCTTGAGCTCCAGAGTTTGAGACCAGCTGGATGAATGCAAAACC	31548
Qy	484	ctgtctctacaaaaataacaaaaatagatgggtggtggcagtcgaccttggctcc	543
Db	31549	CTGTCTCTACAAAAAATAACAAAAATTAGATGGGTGTGGTCATGCACCTGTGTGCTCC	31608
Qy	544	agctacttggaggctaaagtgaggagatcgcttgagccagggagtcaggtctacactg	603
Db	31609	AGCTACTTGGGAGGCTAAGGTGGGAGGATGCTTGTACCCAGGGAGTCAAGTCTACACTG	31668
Qy	604	agccatgattggatcactgcactccagcctcgggttagacagacagaccctgtctcaaaa	663
Db	31669	AGCCATGATTGGATCACTGCATCCAGCTCCAGCTGGGTAGACAGAGAACCCCTGTCTCAAA	31728
Qy	664	aaaagaaatgaagagaaagaaagagagagagagagagagagagagagagagagag	723
Db	31729	AAAAGAAATGAAGAGAGAAAGAAAGAGAGAGAGAGAGAGAGATGAGGGAGGAGGAG	31788
Qy	724	ggggggaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaag	783
Db	31789	GGGGGAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGAT	31848
Qy	784	gaaacagagggcagaagaactttacgttaaatgctcatcatgtgttcaagtgtgacct	843
Db	31849	GAACACAGGCGAGAAAGACTTTACGTAAATTTGCTCATGTGTGTGTGTGTGTGTGTGTGT	31908
Qy	844	caaaacccaattttaccacaggttatctttgactgaggaaggaaggggttcgctctcct	903
Db	31909	CAAAACCCAATTTATTGACCAAGGTTATTCTTTGACTGAGCAAGGGGGTCCGCTCTCCT	31968
Qy	904	gggccttgggctttagaagctcatctctgaccttctctgagatccatccctttcttta	963
Db	31969	GGGCTTGGGCTTTAGAAAGCTCATCTCTGGCCCTTTCTGAGATCCATCCCTTTCTTTTA	32028
Qy	964	ttttcttgacacggagctctgctctgctcactcaggttgagtgagtgagtgagtgagtg	1023
Db	32029	TWTTTCTTGACACGGAGTCTTTGCTCTGTCTACTCAGGCTGGAGTGCAGTGGCATGATCTCG	32088
Qy	1024	actcaacttaacctctgctcccggttcaagcgaatctcctcctcagcctcctcagat	1083
Db	32089	ACTCACTTAACCTCTGCTCTCCCGGTTTCAAGCGGATTTCTCTGCTCAGCCCTCTGAGAT	32148
Qy	1084	aacaggcccccgcaccacatctggttaattttgtatttttagtaagactgggttca	1143









[illegible]

```

|||||
Db 27031 CATCATGTTGCCAGGTTGGTTTCGAACTCTGACCTGAGTGAGTGCCACCTTGGCC 27090
Qy 1202 tcccaagtgctggtgattacagggcatgagccactgcgcccagctccagatccatcccttcc 1261
Db 27091 TCCCAAAAGTGTGGGATTACAGGCATGAGCCACTGTCGCCGCCAGCTCAGATCCATCCCTTTC 27150
Qy 1262 taagggaacacagtcctatggtgcaaaagggccatgcccacccagagttatagttacctggg 1321
Db 27151 TAAGGGCAACAGTCCATGGTGTGAAAGGGGCCATGCCCACCCAGAGTTATAGTACCTGGG 27210
Qy 1322 actccagaattccttgctgctggtgctccacatgcatctccaggccctgcttggtgctct 1381
Db 27211 ACTCCAGAATTCCCTTGCTGGTGGCCCTCCACATGCACTTCCAGGGCTGCTTGGGCTCT 27270
Qy 1382 tctatgctgctgctgctgagtggtgatagaaccactgatgtgagtacctggcttgagcgg 1441
Db 27271 TCTATGGGTCTGTCTGCTGAGTGTGATAGAACCACTGATGTGAGTACCTGGCTTGAGCCG 27330
Qy 1442 tggcctggagatcctgttgactgtagcatgaggggctgtgcatgaatcgaatctgcat 1501
Db 27331 TGGCCTGGAGATCTGTTGACTGTAGCATGAGGGGGCTTGTGCAGCTGAATGCTGXYAT 27390
Qy 1502 gcaagtggtggtggttctggaatatgatgagctgaggtgggaagagaagttaggttgg 1561
Db 27391 GCAGGTGGTGGGAGTTCTGGAATATGATGAGCTGGAGTGGAGAGAAAGTAGGCTGG 27450
Qy 1562 ggcagctctctatgcatccactcttctggccaaaactcaagctcaactgtaagagctcta 1621
Db 27451 GGCAGCTCTCTGAGCCACCTCAFTCTGGCCAAACTCAGGTCAAACTGTGAAGAGCTCA 27510
Qy 1622 aatgtaactgcccctcaagggtggctacaaaggtatcttctgcaaggtagagaccttg 1681
Db 27511 AATGTGAATCTGCCCTTCAAGGTGGCTACAAAGGTATCTTGTCAAGTAGGAGACCTTG 27570
Qy 1682 tggcctccagtgcatctccagggctgctgggctcttctacagggctgctgctgagtc 1741
Db 27571 TGGCCTCCAGTGCATCTCCAGGGCTGCTTGGGCTCTTCTACGGGTCTGTCTCTGAGTC 27630
Qy 1742 ttctatg 1748
Db 27631 TTCTATG 27637

RESULT 7
AC012230
LOCUS Homo sapiens clone RP11-1M10, WORKING DRAFT SEQUENCE, 39 unordered
DEFINITION AC012230 175064 bp DNA linear HTG 22-APR-2000
ACCESSION AC012230
VERSION AC012230.3 GI:7637254
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 175064)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens, clone RP11-1M10
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 175064)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Balwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhalter,B.,
Brown,A., Castile,A., Collangelo,M., Collins,S., Collymore,A.,
Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczy,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,

```

Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,  
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (21-Oct-1999) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Apr 22, 2000 this sequence version replaced gi:6454033.  
All repeats were identified using RepeatMasker:  
Smit,A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information  
Center project name: L2510  
Center clone name: L\_M.10

----- Summary Statistics

Sequencing vector: M13; M7815; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 117571 bases at least Q40  
Consensus quality: 145749 bases at least Q30  
Consensus quality: 160940 bases at least Q20  
Insert size: 185000; agarose-fp  
Insert size: 171264; sum-of-ctnigs  
Quality coverage: 2.9 in Q20 bases; agarose-fp  
Quality coverage: 3.2 in Q20 bases; sum-of-ctnigs

-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 39 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 1003: contig of 1003 bp in length  
\* 1004 1103: gap of 100 bp  
\* 1104 2634: contig of 1531 bp in length  
\* 2635 2734: gap of 100 bp  
\* 2735 4415: contig of 1681 bp in length  
\* 4416 4515: gap of 100 bp  
\* 4516 5785: contig of 1270 bp in length  
\* 5786 5885: gap of 100 bp  
\* 5886 7879: contig of 1994 bp in length  
\* 7880 9686: contig of 1707 bp in length  
\* 9687 9786: gap of 100 bp  
\* 9787 12253: contig of 2467 bp in length  
\* 12254 12353: gap of 100 bp  
\* 12354 15228: contig of 2875 bp in length  
\* 15229 15328: gap of 100 bp  
\* 15329 17200: contig of 1872 bp in length  
\* 17201 17300: gap of 100 bp  
\* 17301 20131: contig of 2831 bp in length  
\* 20132 20231: gap of 100 bp  
\* 20232 22587: contig of 2356 bp in length  
\* 22588 22687: gap of 100 bp  
\* 22688 25707: contig of 3020 bp in length  
\* 25708 25807: gap of 100 bp  
\* 25808 28184: contig of 2377 bp in length  
\* 28185 28284: gap of 100 bp  
\* 28285 31338: contig of 3054 bp in length  
\* 31339 31438: gap of 100 bp  
\* 31439 34299: contig of 2861 bp in length  
\* 34300 34399: gap of 100 bp  
\* 34400 38318: contig of 3919 bp in length  
\* 38319 38418: gap of 100 bp  
\* 38419 42835: contig of 4417 bp in length  
\* 42836 42935: gap of 100 bp  
\* 42936 45448: contig of 2513 bp in length  
\* 45449 45548: gap of 100 bp









Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
 Howland, J.C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,  
 Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G., Locke, K.,  
 Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,  
 McPheeters, R., Meldrum, J., Meneus, L., Morrow, J., Naylor, J.,  
 Norman, C.H., O'Connor, T., O'Donnell, P., Oliver, T.M., Peterson, K.,  
 Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,  
 Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,  
 Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
 Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W.J.,  
 Zimmer, A. and Zody, M.

TITLE  
 JOURNAL  
 COMMENT

Direct Submission  
 Submitted (18-JAN-2000) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Jul 13, 2000 this sequence version replaced gi:6705871.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: <http://www-seq.wi.mit.edu>  
 Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
 ----- Project Information  
 Center project name: L2512  
 Center clone name: L\_N10  
 -----

\* NOTE: This record contains 73 individual  
 \* sequencing reads that have not been assembled into  
 \* contigs. Runs of N are used to separate the reads  
 \* and the order in which they appear is completely  
 \* arbitrary. Low-pass sequence sampling is useful for  
 \* identifying clones that may be gene-rich and allows  
 \* overlap relationships among clones to be deduced.  
 \* However, it should not be assumed that this clone  
 \* will be sequenced to completion. In the event that  
 \* the record is updated, the accession number will  
 \* be preserved.

1  
 \* 872 971: contig of 871 bp in length  
 \* 972 1834: contig of 863 bp in length  
 \* 1835 1934: gap of 100 bp  
 \* 1935 2804: contig of 870 bp in length  
 \* 2805 2904: gap of 100 bp  
 \* 2905 3745: contig of 841 bp in length  
 \* 3746 3845: gap of 100 bp  
 \* 3846 4696: contig of 851 bp in length  
 \* 4697 4796: gap of 100 bp  
 \* 4797 5640: contig of 844 bp in length  
 \* 5641 5740: gap of 100 bp  
 \* 5741 6540: contig of 800 bp in length  
 \* 6541 6640: gap of 100 bp  
 \* 6641 7509: contig of 869 bp in length  
 \* 7510 7609: gap of 100 bp  
 \* 7610 8479: contig of 870 bp in length  
 \* 8480 8579: gap of 100 bp  
 \* 8580 9430: contig of 851 bp in length  
 \* 9431 9530: gap of 100 bp  
 \* 9531 10376: contig of 846 bp in length  
 \* 10377 10476: gap of 100 bp  
 \* 10477 11322: contig of 846 bp in length  
 \* 11323 11422: gap of 100 bp  
 \* 11423 12302: contig of 880 bp in length  
 \* 12303 12402: gap of 100 bp  
 \* 12403 13280: contig of 878 bp in length  
 \* 13281 13380: gap of 100 bp  
 \* 13381 14241: contig of 861 bp in length  
 \* 14242 14341: gap of 100 bp  
 \* 14342 15196: contig of 855 bp in length  
 \* 15197 15296: gap of 100 bp  
 \* 15297 16123: contig of 827 bp in length  
 \* 16124 16223: gap of 100 bp  
 \* 16224 17072: contig of 849 bp in length

17073 17172: gap of 100 bp  
 \* 17173 18041: contig of 869 bp in length  
 \* 18042 18141: gap of 100 bp  
 \* 18142 19009: contig of 868 bp in length  
 \* 19010 19109: gap of 100 bp  
 \* 19110 19966: contig of 857 bp in length  
 \* 19967 20066: gap of 100 bp  
 \* 20067 20921: contig of 855 bp in length  
 \* 20922 21021: gap of 100 bp  
 \* 21022 21865: contig of 844 bp in length  
 \* 21866 21965: gap of 100 bp  
 \* 21966 22832: contig of 867 bp in length  
 \* 22833 22932: gap of 100 bp  
 \* 22933 23780: contig of 848 bp in length  
 \* 23781 23880: gap of 100 bp  
 \* 23881 24733: contig of 853 bp in length  
 \* 24734 24833: gap of 100 bp  
 \* 24834 25670: contig of 837 bp in length  
 \* 25671 25770: gap of 100 bp  
 \* 25771 26621: contig of 851 bp in length  
 \* 26622 26721: gap of 100 bp  
 \* 26722 27576: contig of 855 bp in length  
 \* 27577 27676: gap of 100 bp  
 \* 27677 28532: contig of 856 bp in length  
 \* 28533 28632: gap of 100 bp  
 \* 28633 29492: contig of 860 bp in length  
 \* 29493 29592: gap of 100 bp  
 \* 29593 30455: contig of 863 bp in length  
 \* 30456 30555: gap of 100 bp  
 \* 30556 31410: contig of 855 bp in length  
 \* 31411 31510: gap of 100 bp  
 \* 31511 32368: contig of 858 bp in length  
 \* 32369 32468: gap of 100 bp  
 \* 32469 33312: contig of 844 bp in length  
 \* 33313 33412: gap of 100 bp  
 \* 33413 34268: contig of 856 bp in length  
 \* 34269 34368: gap of 100 bp  
 \* 34369 35204: contig of 836 bp in length  
 \* 35205 35304: gap of 100 bp  
 \* 35305 36156: contig of 852 bp in length  
 \* 36157 36256: gap of 100 bp  
 \* 36257 37128: contig of 872 bp in length  
 \* 37129 37228: gap of 100 bp  
 \* 37229 38083: contig of 855 bp in length  
 \* 38084 38183: gap of 100 bp  
 \* 38184 39031: contig of 848 bp in length  
 \* 39032 39131: gap of 100 bp  
 \* 39132 40006: contig of 875 bp in length  
 \* 40007 40106: gap of 100 bp  
 \* 40107 40967: contig of 861 bp in length  
 \* 40968 41067: gap of 100 bp  
 \* 41068 41913: contig of 846 bp in length  
 \* 41914 42013: gap of 100 bp  
 \* 42014 42824: contig of 811 bp in length  
 \* 42825 42924: gap of 100 bp  
 \* 42925 43776: contig of 852 bp in length  
 \* 43777 43876: gap of 100 bp  
 \* 43877 44752: contig of 876 bp in length  
 \* 44753 44852: gap of 100 bp  
 \* 44853 45724: contig of 872 bp in length  
 \* 45725 45824: gap of 100 bp  
 \* 45825 46643: contig of 819 bp in length  
 \* 46644 46743: gap of 100 bp  
 \* 46744 47599: contig of 856 bp in length  
 \* 47600 47699: gap of 100 bp  
 \* 47700 48551: contig of 852 bp in length  
 \* 48552 48651: gap of 100 bp  
 \* 48652 49485: contig of 834 bp in length  
 \* 49486 49585: gap of 100 bp  
 \* 49586 50440: contig of 855 bp in length  
 \* 50441 50540: gap of 100 bp  
 \* 50541 51404: contig of 864 bp in length  
 \* 51405 51504: gap of 100 bp



```

* 51505 52372: contig of 868 bp in length
* 52373 52472: gap of 100 bp
* 52473 53328: contig of 856 bp in length
* 53329 53428: gap of 100 bp
* 53429 54268: contig of 840 bp in length
* 54269 54368: gap of 100 bp
* 54369 55229: contig of 861 bp in length
* 55230 55329: gap of 100 bp
* 55330 56197: contig of 868 bp in length
* 56198 56297: gap of 100 bp
* 56298 57163: contig of 866 bp in length
* 57164 57263: gap of 100 bp
* 57264 58130: contig of 867 bp in length
* 58131 58230: gap of 100 bp
* 58231 59082: contig of 852 bp in length
* 59083 59182: gap of 100 bp
* 59183 60020: contig of 838 bp in length
* 60021 60120: gap of 100 bp
* 60121 60983: contig of 863 bp in length
* 60984 61083: gap of 100 bp
* 61084 61935: contig of 852 bp in length
* 61936 62035: gap of 100 bp
* 62036 62866: contig of 831 bp in length
* 62867 62966: gap of 100 bp
* 62967 63827: contig of 861 bp in length
* 63828 63927: gap of 100 bp
* 63928 64783: contig of 856 bp in length
* 64784 64883: gap of 100 bp
* 64884 65740: contig of 857 bp in length
* 65741 65840: gap of 100 bp
* 65841 66684: contig of 844 bp in length
* 66685 66784: gap of 100 bp
* 66785 67651: contig of 867 bp in length
* 67652 67751: gap of 100 bp

Query Match          10.8%; Score 348; DB 2; Length 69570;
Best Local Similarity 99.7%; Pred. No. 7.2e-190;
Matches 398; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2696 atgtgtcgtggcggtgaaagtcgcccgtttaagggcgggcccccgcctccacgtgctt 2755
Db 41367 ATGTGTCTGTGGCGGCTGAACGTGCGCGGTAAAGGGCGGGCCCGGCTCCACGTGCTT 41426

Qy 2756 tctactaagtactgaactacataaacagagccgagcggcgagggcgggagggagagag 2815
Db 41427 TCTGTAGTGACTGAACACTACTATAACAGAGCCCGGACAGCGGGCGGGAGGGAGAG 41486

Qy 2816 cacaggctttgaccgatagtagtaacctcgtcgtcgtgtagcagccgaatctataaaaggaact 2875
Db 41487 CACAGGCTTTGACCGATAGTAGTAACCTCTGCGTCTGTCAGCCGAATCTATAAAAGGAAC 41546

Qy 2876 agtcccgcaaaaaccccgtaattgcgagcagagagtagtgaggcgccgggaccgcagagc 2935
Db 41547 AGTCCCGCAAAAACCCCGTAATTTCGAGCAGAGAGTAGTGAGTGGGGCCGGACCCCGCAGAGC 41606

Qy 2936 cgagccgaccttctctccgggctcggcagggcagggcgagctccgcgcacacac 2995
Db 41607 CGAGCCGACCTTCTCTCCGGGCTGCGCAGGCGGCGGAGGCTCCGCGCACCAAC 41666

Qy 2996 agagcggattctcagggcgcttctcctctgttttttcccccggttctgttttctccctt 3055
Db 41667 AGAGCGGCTTCTCAGGCGGCTTGTCTCTCTGTTTTTCCCGGTTCTGTTTTCTCCGCTT 41726

Qy 3056 ctccggaaggcttgcaagggttaggaagaagacaca 3094
Db 41727 CTCGGAAGGCTTGTCAGAGGGGTAGGAGAAAGAGACGCA 41765

RESULT 13
AC012230/c
LOCUS
DEFINITION Homo sapiens clone RP11-1M10, WORKING DRAFT SEQUENCE, 39 unordered
pieces.

```

# ACCESSION VERSION KEYWORDS SOURCE ORGANISM

AC012230  
AC012230.3 GI:7637254  
HTGS\_PHASE1; HTGS\_DRAFT.  
human.

## REFERENCE

1 (bases 1 to 175064)  
Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
Homo sapiens, clone RP11-1M10  
Unpublished  
2 (bases 1 to 175064)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,  
Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhgalter,B.,  
Brown,A., Castelle,A., Colangelo,M., Collins,S., Collamore,A.,  
Cooke,P., DeArellano,K., Dewar,K., Domino,M., Donegan,L., Doyle,M.,  
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,  
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Hearford,A., Horton,L.,  
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,  
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,  
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,  
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,  
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,  
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

## TITLE

JOURNAL

## REFERENCE

AUTHORS

AUTHORS

JOURNAL

COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L2510

Center clone name: L2510

----- Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 117571 bases at least Q40

Consensus quality: 145749 bases at least Q30

Consensus quality: 160940 bases at least Q20

Insert size: 185000; agarose-fp

Insert size: 171264; sum-of-contigs

Quality coverage: 2.9 in Q20 bases; agarose-fp

Quality coverage: 3.2 in Q20 bases; sum-of-contigs

-----

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 39 contigs. The true order of the pieces

\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will

\* be preserved.

\* 1 1003: contig of 1003 bp in length

\* 1004 1103: gap of 100 bp

\* 1104 2634: contig of 1531 bp in length

\* 2635 2734: gap of 100 bp

\* 2735 4415: contig of 1681 bp in length

\* 4416 4515: gap of 100 bp

\* 4516 5785: contig of 1270 bp in length

\* 5786 5885: gap of 100 bp

\* 5886 7879: contig of 1994 bp in length

\* 7880 7979: gap of 100 bp

\* 7980 9686: contig of 1707 bp in length

\* 9687 9786: gap of 100 bp



```

|||||
Db 1405 TCCCAAGGAGCTCAAGAGTTTCAACAAAGTTTCACTTTTCAGAAACCCCTTTGAGGAGAC 1346
|||||
Qy 335 agaataacatctctctccattttaaagatgaagaacagggccgggacacattggctaat 394
|||||
Db 1345 AGAATATACATCTCTCTCCATTTTAAAGATGAAGAAACAGCGCGGCACAAATGGCTAAT 1286
|||||
Qy 395 gctgttaatccacacatttggagggtgagcca 429
|||||
Db 1285 GCCTGTATCCAGCACTTTGGGAGGCTGAGGCA 1251
|||||

RESULT 14
AK022254 1750 bp mRNA linear PRI 29-SEP-2000
LOCUS Homo sapiens cDNA FLJ12192 fis, clone MAMMAL000851.
ACCESSION AK022254
VERSION AK022254.1 GI:10433612
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens Mammary gland cDNA to mRNA, clone_lib:MAMMAL
clone:MAMMAL000851.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Takahashi-Fujii,A., Hara,H.,
Tanase,T., Nomura,Y., Todaya,S., Komai,F., Hara,R., Takeuchi,K.,
Arita,M., Nabekura,T., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J.,
Wakamatsu,A., Nakamura,Y., Nagahara,K., Masuho,Y. and Oshima,A.
NEDO human cDNA sequencing project
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished (2000)
REFERENCE 2 (bases 1 to 1750)
AUTHORS Isogai,T. and Otsuki,T.
DIRECT SUBMISSION
TITLE Direct Submission
JOURNAL Submitted (23-AUG-2000) to the DDBJ/EMBL/GenBank databases. Takao
Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana,
Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@helix.co.jp,
Tel:81-438-52-3951, Fax:81-438-52-3952)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing and clone selection;
Helix Research Institute (supported by Japan Key Technology Center
etc.) and Department of Virology, Institute of Medical Science,
University of Tokyo.
FEATURES
Location/Qualifiers
source
1..1750
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MAMMAL000851"
/clone_lib="MAMMAL"
/tissue_type="Mammary gland"
/note="cloning vector: pME18SFL3"
BASE COUNT 291 a 489 c 586 g 384 t
ORIGIN

Query Match 10.4%; Score 336; DB 9; Length 1750;
Best Local Similarity 100.0%; Pred. No. 4.6e-183;
Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2896 aattcgagcagagtgagtgaggccgagaccgagagccgagagccctctctccc 2955
|||||
Db 1 AATTGCGAGCAGAGTGAGTGGGCGGGACCCGCGAGAGCCGAGCCGCTTCTCTCCC 60
|||||
Qy 2956 gggctgcgcagggcagggcgaggctccgcgcaccacacagagccggttctcaggggc 3015
|||||
Db 61 GGGCTGCGCGAGCGGCGGGGAGCTCCGGCACCACAGAGCGGTTCTCAGGGCGC 120
|||||
Qy 3016 ttgtctctgtttttcccccgttctgttttctccctctccgaaggttgcagg 3075
|||||
Db 121 TTTGTCTCTGTTTTCCTCCCGGTTCTGTTTCTCCCTTCTCCGGAAGGTTGTCAAGG 180
|||||

```

```

Qy 3076 ggtaggagaaagagagcgcaaacacacaaagtgaacaggtctccagtact 3135
|||||
Db 181 GGTAGGAGAAAGAGAGCGCAACACAAAAGTGGAAAACAGGTAAAGAGCTCTCCAGTGACT 240
|||||
Qy 3136 tacttggcggtattatttttggaggcccaaggaggtctcggaagtgcctcggttcg 3195
|||||
Db 241 TACTTGGCGGTATTGTTTGTTCGAGGCCAAGGAGGCTTCGGGAAGTGCTCGGTTTCG 300
|||||
Qy 3196 gggactttgatccggagccccacatccccaccactt 3231
|||||
Db 301 GGGACTTTGATCGGAGCCCATCCACCACCTT 336
|||||

RESULT 15
AF258627 697 bp mRNA linear PRI 11-MAY-2000
LOCUS Homo sapiens ATP binding cassette transporter 1 (ABCA1) mRNA,
partial cds.
ACCESSION AF258627
VERSION AF258627.1 GI:7769707
KEYWORDS human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 697)
AUTHORS Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C.,
Aouizerat,B.E., Fielding,C.J. and Kane,J.P.
ANALYSIS OF hABCI gene 5' end: additional peptide sequence,
promoter region, and four polymorphisms
Biochem. Biophys. Res. Commun. 271 (2000) In press
JOURNAL 2 (bases 1 to 697)
REFERENCE Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C.,
Aouizerat,B.E., Fielding,C.J. and Kane,J.P.
DIRECT SUBMISSION
TITLE Direct Submission
JOURNAL Submitted (19-APR-2000) Cardiovascular Research Institute,
University of California, San Francisco, 505 Parnassus Avenue, San
Francisco, CA 94143-0130, USA
FEATURES
Location/Qualifiers
source
1..697
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/map="9q31"
/tissue_type="placenta"
1..>697
/gene="ABCA1"
/note="ABC1"
396..>697
/gene="ABCA1"
/note="membrane-bound"
/codon_start=1
/product="ATP binding cassette transporter 1"
/protein_id="AAF69513.1"
/db_xref="GI:7769708"
/translation="MACWFLRLLLKMLNLTERRRQTLLELVANPLFLILISVRL
SYPPVEQHECHPKNKAPSAGTLPWVGIIICNANNPCFRYPPTGPEAGVVGNNKS"
BASE COUNT 152 a 198 c 190 g 156 t
ORIGIN

Query Match 9.4%; Score 303; DB 9; Length 697;
Best Local Similarity 100.0%; Pred. No. 6.1e-164;
Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2812 agagcacaggtttgaccgatagtaacctctgcctcggtgcagccgaatctataaagg 2871
|||||
Db 1 AGACACAGAGGCTTTGACCGTAGTAACCTCTCGCTCGTCCGTCAGCCGAACTATATAAAGG 60
|||||
Qy 2872 aactgacctccggcaaaaccccgctaaattgcgagcgagtgagtgaggccggaccgcga 2931
|||||
Db 61 AACTAGTCCCGGCAAAACCCCGTAATTGCGAGCGAGTGAGTGGGGCCGGGACCGCA 120
|||||

```







diagnostic application  
Patent: WO 0130848-A 70 03-MAY-2001;  
Aventis Pharma S.A. (FR)  
Location/Qualifiers  
source 1. .9854

BASE COUNT	2665 a	2219 c	2334 g	2635 t	1 others
ORIGIN					

Query Match 6.3%; Score 205; DB 6; Length 9854;  
Best Local Similarity 100.0%; Pred. No. 6.7e-107;  
Matches 205; Conservative 0; Mismatches 0; Indels

Qy	2910	gtagtggggcgagaccgcagagccagaccctctctccgggctcgagcaagg	2969
Db	1	GTGAGTGGGCGCGGACCGCGAGCGAGCGAGCCCTCTCTCCGGGCTGCGCGCAGG	60
Qy	2970	caaggcgaggagctccgcgcacacacagcggcttcagggcgcttgcctctgtt	3029
Db	61	CAGGGCGGGGAGCTCCGCGCGCCACAGAGCCGGTTCAGGGCGCTTTGCTCCTGTTT	120
Qy	3030	tttcccgcggtctctgtttctcccttctccggaagcgcttgcaggggtagagagaaag	3089
Db	121	TTTCCCGGGTCTGTGTTTTCTCCCTCTCCGGAAGGCTGTCAAGGGGTAGAGAAAGAG	180
Qy	3090	acgcacacacaaaagtggaaaaag	3114
Db	181	ACGCAACACAAAAGTGGAAACAG	205

RESULT	21	
AX139818		
LOCUS	AX139818	9854 bp
DEFINITION	Sequence 70 from Patent EP1096012.	DNA
ACCESSION	AX139818	linear
VERSION	AX139818.1	
	GI:14275400	
		PAT 30-MAY-2001

REINFORCED SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 9854) Denefle, P., Rosier-Montus, M.F., Arnould-Reguigne, I., Prades, C., Naudin, L., Lemoine, C., Duverger, N., Jaye, M., searfooss Iil, G.H., Remaley, A., Brewer, H.B. and Dean, M.
TITLE	Nucleic acids of the human abcl gene and their therapeutic and diagnostic application
JOURNAL	Patent: Ep 1096012-A 70 02-MAY-2001:

```

source
1. .9854
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2665 a 2219 c 2334 g 2635 t
ORIGIN
1 others

```

Query Match 6.3%; Score 205; DB 6; Length 9854;  
Best Local Similarity 100.0%; Pred. No. 6.7e-107;  
Matches 205; Conserving 0; Mismatches 0; Indels 0; Gaps 0;

QY	2910	gtgagtggggcccggagaccgcagagccgagccgagccacctctctcccggggtgcggcaggg	2965
Db	1	GTGAGTGGGGCCGGGACCCGCGAGAGCCGAGAGCCGACCTTCTCTCCGGGCTCGCGCAGGG	60
QY	2970	cagggcggggagctccgcgcacaaacagagccggttctcaggcgcttgctctgttt	3029
Db	61	CAGGCGGGGAGCTCCCGCGCACCAAGAGCCGGTTCTCAGGCGCTTTGGCTTCCTGTGTT	120
QY	3030	tttcccggttctgtttcttccctctctccggaagcttgtaaaagggttaggagaaaagag	3089

Db 121 TTTCCCGGTTCTGTTTTTCTCCCTTCTCCGGAAGGCTTGTCAAGGGTAGGAGAAAGAG 180

Qy 3090 acgcaaacacacaaaagtggaaaaacag 3114

Db 181 ACGCAAAACACAAAAGTGGAAACAG 205

RESULT	22	DNA	linear	PAT 22-JAN-2001
AX060713	AX060713	10442 bp		
LOCUS	AX060713			
DEFINITION	Sequence 1 from Patent WO0078972.			
ACCESSION	AX060713			
VERSION	AX060713.1	GI:12406103		
KEYWORDS				
SOURCE	human			

**ORGANISM** Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
**REFERENCE** 1 (bases 1 to 10442)  
**AUTHORS** Lawn, R.M., Wade, D. and Garvin, M.  
**TITLE** Regulation with binding cassette transporter protein abcl  
**JOURNAL** Patent: WO 0078972-A 1 28-DEC-2000;

```

source
I. .10442
/organism="Homo sapiens"
/db_xref="taxon:9606"
2898 a 2297 c 2408 g 2835 t
BASE COUNT
ORIGIN
4 others

```

Query Match 6.1%; Score 197; DB 6; Length 10442;  
Best Local Similarity 100.0%; Pred. NO. 3e-102;

Qy	2918	ggcgggaccgcgcagagccgacccctctctccggctgcgagggcagggcgg	2977
Db	1	GGCGGGACCGAGAGCCGAGCCGAGCCGACCTTCTCTCCGGGTGGCGAGGCGAGGCGG	60
Qy	2978	ggaactcgcgcacccaacagacgggtctcagggcggttgcctctgtttttcccg	3037
Db	61	GGAGCTCCGGGCACCAACAGAGCGGGTCTCAGGGCGCTTGTCTCTGTCTTTTTCGCCG	120
Qy	3038	gtctgtttttccctctcccggaaggcttgtcaagggttaggagaaagagacgcaaac	3097
Db	121	GTTCTGTCTTTCTCCCTTCTCCGAGGGCTGTCTAAGGGTAGGAAAGACGCAAC	180
Qy	3098	acaaagtggaaaaacag	3114
Db	181	ACAAAGTGGAAACAG	197

RESULT	23
AX060892	
LOCUS	AX060892
DEFINITION	Sequence 1 from Patent WO0078971.
ACCESSION	AX060892
VERSION	AX060892.1 GI:12406270
	linear DNA 10442 bp
	PAT 22-JAN-2001

SOURCE: Homo sapiens  
ORGANISM: Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE: 1 (bases 1 to 10442)  
AUTHORS: Lawn, R.M., Wade, D., Oram, J.F. and Garvin, M.  
TITLE: ATP binding cassette transporter protein abcl polypeptides  
JOURNAL: Patent: WO 0078971-A 1 28-DEC-2000.

	source	i .10442	/organism="Homo sapiens"	
			/db_xref="taxon:9606"	
BASE COUNT		2898 a	2297 c	2408 q 2835 t 4 others

## ORIGIN

Query Match 6.1%; Score 197; DB 6; Length 10442;  
 Best Local Similarity 100.0%; Pred. No. 3e-102;  
 Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2918 ggcgggagccgagagccgagccgagccctctctccggcgctgagcgagcgagcgcg 2977  
 |||||  
 Db 1 ggcgggagccgagagccgagccgagccctctctccggcgctgagcgagcgagcgcg 60  
 |||||

Qy 2978 ggaactcgcgacacacagagcggtctcaggggcgcttctcctgtttttccccc 3037  
 |||||  
 Db 61 ggagctcgcgacacacagagcggtctcaggggcgcttctcctgtttttccccc 120  
 |||||

Qy 3038 gttctgtttctccctctccgagagcttctcaagggttagagaaagagcgcaaac 3097  
 |||||  
 Db 121 gttctgtttctccctctccgagagcttctcaagggttagagaaagagcgcaaac 180  
 |||||

Qy 3098 acaaaagtggaaacag 3114  
 |||||  
 Db 181 acaaaagtggaaacag 197  
 |||||

## RESULT 24

AF285167  
 LOCUS  
 DEFINITION  
 Homo sapiens ATP-binding cassette transporter 1 (ABCA1) mRNA,  
 complete cds.  
 ACCESSION  
 AF285167  
 VERSION  
 AF285167.1 GI:9755158  
 KEYWORDS  
 human.  
 SOURCE  
 human.  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1 (bases 1 to 10442)  
 AUTHORS  
 Schwartz, K., Lawn, R.M. and Wade, D.P.  
 TITLE  
 ABCA1 gene expression and apoA-I-mediated cholesterol efflux are  
 regulated by LXR  
 JOURNAL  
 Unpublished  
 REFERENCE  
 2 (bases 1 to 10442)  
 AUTHORS  
 Lawn, R.M., Wade, D.P., Garvin, M.R., Wang, X., Schwartz, K.,  
 Porter, J.G., Seilhamer, J.J., Vaughan, A.M. and Oram, J.F.  
 TITLE  
 Direct Submission  
 JOURNAL  
 Submitted (06-JUL-2000) Discovery Research, CV Therapeutics Inc.,  
 3172 Porter Drive, Palo Alto, CA 94304, USA  
 FEATURES  
 Location/Qualifiers  
 1..10442  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="9"  
 /map="9q31"  
 /cell\_type="fibroblast"  
 /tissue\_type="skin"  
 1..10442  
 /gene="ABCA1"  
 291..7076  
 /gene="ABCA1"  
 /codon\_start=1  
 /product="ATP-binding cassette transporter 1"  
 /protein\_id="AAF98175.1"  
 /db\_xref="GI:9755159"  
 /translation="MACWPQLRLLLKNNLTPRRROTQQLLEVAWPLIFILISVRL  
 SYPPYEHECHFPNKPMSAGTLPMWQICINANNPCFRYPPTGEAFGVGNFNKSIY  
 ARLFSDARLLLYSKQDTSMDKMRKVLRTLOQIKKSSNKLQDFLDVNDTEFSGLYH  
 NLSLPKSVDRVLRNMDILKTLRLNSTSPPPSKELAEAKTLHLSTLAQELFSMR  
 KEKLAARQEVMLTWNSSSTQYQAVSRVCGHPGEGGLKIKSLNWEEDNNTKALF  
 SWSMDRQEVMLTWNSSSTQYQAVSRVCGHPGEGGLKIKSLNWEEDNNTKALF  
 GGNGTEADAEFYDNSTPYPCNDLMKNLESSPLRIWKALKPLVLYKILYTPDTPAT  
 QQWAEVNTQELAVFADHLEGMEELSPKIVTFMENSQMDVLRMLDSRDNDHWE  
 QDLGLDWTQADIVAFIAKHPDEVOSSNGSVYITREFNETNQAIRTSRFMECVNLN  
 KLEPIATEVWLINKSMELLDERKFWAGIVFTGTPGSIELPHHVKKYKIRMDIDNVERT

## gene

## CDS

RESULT 25  
 AX060719  
 LOCUS  
 DEFINITION  
 Sequence 7 from Patent WO0078972.  
 ACCESSION  
 AX060719  
 VERSION  
 AX060719.1 GI:12406108  
 KEYWORDS  
 SOURCE  
 human.  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1 (bases 1 to 10474)  
 AUTHORS  
 Lawn, R.M., Wade, D. and Garvin, M.  
 TITLE  
 Regulation with binding cassette transporter protein abcl  
 JOURNAL  
 Patent: WO 0078972-A 7 28-DEC-2000;  
 CV THERAPEUTICS, INC. (US)  
 FEATURES  
 Location/Qualifiers  
 1..10474  
 10474 bp  
 DNA  
 linear  
 PAT 22-JAN-2001

NKIKDGYWDPGRADPFEDMBYVMGGEAYLODVVEQAIIRVLTGTEKTKGYVMQMPY  
 PCYVDDIFLRVWSSMPLFMTLAWIISVAVIIGIKGIVIEKEARLKTMRIMGDNSILW  
 FSWFISSLIPLLLVAGLLVILKLGKLVNPPSDPVVVFVSVFVAVTILQCFLLSTLF  
 SRNLAAACGGIITFTLYLVYLCVAMQDYVFTLKIFASLLSPVAFGCGEYFALFE  
 EQGIGVQNDLNFESDEBDFNLTTISMLFDTFLXGVTWYIEAVPPGOGYIPRPW  
 YFPCTKSYWFGESDEKSHPGSNOKRSEICMEEPETHLKLGVSIQNLVKYVRDGMKV  
 AVDLGLAVYEQITSTFLHGNGAKTTTMSILTLGLFPTSGTAVILCKDISESTIR  
 ONLGVCPHNVILFOMLVVEEHIWYARLKLSEKHVAKEQEMALDVLGSPSKLSKT  
 SOLSGMORKLSVALAFVGGSVILDEPTAGVDPIYSRRIWELLATROGRTILIST  
 HMDADVLDGRIAIISHGKLCVCGSSLFKLNQLTGYITLVYLVKKDVSSLSRGNLS  
 SYVYLKEDSVSSOSSDAGLSDHESDTLTIDVSAISNLIRKHSARLIVSDESHGEL  
 TYVLPYAAKEGAFVLFHEIDRDLSDLGISSETTLEIFLKVAESEGVDAETS  
 DGLTPARNRRAFGDKSLRPFETDDAADPNDSDIDPESRETLLSGMDGKGSYQVK  
 GWKLTQOQFVALLMKRLLIARSRKGFQIIVLPAVFCIALVPSLIVPPGKYPSE  
 LQPMWNEQYTFVSNADPEDTGTLELNLALNLTGFCGTRCEGPNIPDTPQAGEEM  
 TAPVPTIMOLFQNGNMTQNPQCCSSDKIKMLPVCPPGAGGLPPPPQKNTA  
 DILQDLTGRNLSYLVYQIYIAKSLNKIWNVEYFGFSLGVSNQALPPSQEYN  
 DAIOKMKHLKLAKDSSADRLNSLGRFMTGLDRNNVKVFNKNGHAISSFLNIN  
 NAILRANLQENPSHYGITAFAHPLNLTQOOLSEVALMTTSVDVLVSIQVIFAMSV  
 PASFVFLIOERSKAKHLOFTSGVKPVIVLNSFVNDMNVYVPATLVIITFCFOQ  
 KSYVSNLPLVALILLYCHSITPLMPASFPVKIESTAYVLTSLNLTGINSVA  
 TFVLEFTDNKNINILKSVFLIFPHFCGLRGILDMVKNQAMADALERFENRFVS  
 PLSWDLVGRNLFAMAVEGVFEFLITVLIQYRFFIRPRVNAKLSLPDDEDEDVRRERQ  
 RILDGGQNDILEIKELTKIYRRKRKPAVDRCVGPPEGCGLLGVNGAGKSTFKM  
 LTGDTVTTRGDAFLNKSILNIHEVQNMGYCQFDATITELLTGRHVBFFALLRGV  
 PEKVKGVGEWAIRKGLVYGEVAGNYSGNKRKLSTAMALIGGPPVFLDEPTTG  
 MDPKARFLGNALSVMKESVSVLTSHSMECEALCTRMALVNGFRCLGSHVHLK  
 NRGDGTIVVRIAGSNPDLPQVDFGLAFPPGSVLAKHKNMLQYOLPSSLSLARI  
 FSLISQSKRLHIEDIYSVSTQLDQVFNFAKQSDDDDLKDLJSLHKNQTVVDVAVLT  
 SFLQDEKVKESYV

BASE COUNT 2898 a 2297 c 2408 g 2835 t 4 others  
 ORIGIN

Query Match 6.1%; Score 197; DB 9; Length 10442;  
 Best Local Similarity 100.0%; Pred. No. 3e-102;  
 Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2918 ggcgggagccgagagccgagccctctctccggcgctgagcgagcgagcgcg 2977  
 |||||  
 Db 1 ggcgggagccgagagccgagccctctctccggcgctgagcgagcgagcgcg 60  
 |||||

Qy 2978 ggaactcgcgacacacagagcggtctcaggggcgcttctcctgtttttccccc 3037  
 |||||  
 Db 61 ggagctcgcgacacacagagcggtctcaggggcgcttctcctgtttttccccc 120  
 |||||

Qy 3038 gttctgtttctccctctccgagagcttctcaagggttagagaaagagcgcaaac 3097  
 |||||  
 Db 121 gttctgtttctccctctccgagagcttctcaagggttagagaaagagcgcaaac 180  
 |||||

Qy 3098 acaaaagtggaaacag 3114  
 |||||  
 Db 181 acaaaagtggaaacag 197  
 |||||

## RESULT 25

## AX060719

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## FEATURES

## source



```
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2906 a 2305 c 2416 g 2843 t 4 others
ORIGIN

Query Match 5.8%; Score 188; DB 6; Length 10474;
Best Local Similarity 100.0%; Pred. No. 5e-97;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2927 ccgcagagccgagccaccttctctccgggctgcggcagggcagggcgagctccg 2986
Db 42 CCGCAGAGCCGAGCCGACCCCTTCTCCCGGGCTGCGGACGGCAGGGCGGGAGCTCCG 101
QY 2987 cgcaccaacagagccggttctcagggcgcttgcctctgtttttcccggtctgttt 3046
Db 102 CGCACCAACAGAGCCGGTTCTCAGGGCGCTTGTCTCTGTTTTTCCCGGGTCTGTTT 161
QY 3047 tctcccttctccgaaaggcttgcagggttaggagaaagagacgcaacacaaagtg 3106
Db 162 TCTCCCTTCTCCGGAAGGCTTGTCAAGGGTAGGAGAAAGAGACGCAACAAAGTG 221
QY 3107 gaaaacag 3114
Db 222 GAAACAG 229

RESULT 26
AX060721
LOCUS AX060721 10474 bp DNA linear PAT 22-JAN-2001
DEFINITION Sequence 9 from Patent WO0078972.
ACCESSION AX060721
VERSION AX060721.1 GI:12406109
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Lawn,R.M., Wade,D., and Garvin,M.
JOURNAL Regulation with binding cassette transporter protein abcl
PATENT: WO 0078972-A 9 28-DEC-2000;
CV THERAPEUTICS, INC. (US)
FEATURES
source
Location/Qualifiers
1..10474
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2907 a 2304 c 2415 g 2844 t 4 others
ORIGIN

Query Match 5.8%; Score 188; DB 6; Length 10474;
Best Local Similarity 100.0%; Pred. No. 5e-97;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2927 ccgcagagccgagccaccttctctccgggctgcggcagggcagggcgagctccg 2986
Db 42 CCGCAGAGCCGAGCCGACCCCTTCTCCCGGGCTGCGGACGGCAGGGCGGGAGCTCCG 101
QY 2987 cgcaccaacagagccggttctcagggcgcttgcctctgtttttcccggtctgttt 3046
Db 102 CGCACCAACAGAGCCGGTTCTCAGGGCGCTTGTCTCTGTTTTTCCCGGGTCTGTTT 161
QY 3047 tctcccttctccgaaaggcttgcagggttaggagaaagagacgcaacacaaagtg 3106
Db 162 TCTCCCTTCTCCGGAAGGCTTGTCAAGGGTAGGAGAAAGAGACGCAACAAAGTG 221
QY 3107 gaaaacag 3114
Db 222 GAAACAG 229

RESULT 27
AX060898
LOCUS AX060898 10474 bp DNA linear PAT 22-JAN-2001
DEFINITION Sequence 7 from Patent WO0078971.
ACCESSION AX060898
VERSION AX060898.1 GI:12406275
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Lawn,R.M., Wade,D., Oram,J.F. and Garvin,M.
JOURNAL ATP binding cassette transporter protein abcl polypeptides
PATENT: WO 0078971-A 7 28-DEC-2000;
CV THERAPEUTICS, INC. (US)
FEATURES
source
Location/Qualifiers
1..10474
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2907 a 2304 c 2415 g 2844 t 4 others
ORIGIN

Query Match 5.8%; Score 188; DB 6; Length 10474;
Best Local Similarity 100.0%; Pred. No. 5e-97;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2927 ccgcagagccgagccaccttctctccgggctgcggcagggcagggcgagctccg 2986
Db 42 CCGCAGAGCCGAGCCGACCCCTTCTCCCGGGCTGCGGACGGCAGGGCGGGAGCTCCG 101
QY 2987 cgcaccaacagagccggttctcagggcgcttgcctctgtttttcccggtctgttt 3046
Db 102 CGCACCAACAGAGCCGGTTCTCAGGGCGCTTGTCTCTGTTTTTCCCGGGTCTGTTT 161
QY 3047 tctcccttctccgaaaggcttgcagggttaggagaaagagacgcaacacaaagtg 3106
Db 162 TCTCCCTTCTCCGGAAGGCTTGTCAAGGGTAGGAGAAAGAGACGCAACAAAGTG 221
QY 3107 gaaaacag 3114
Db 222 GAAACAG 229
```



\* 34536 35433: contig of 898 bp in length  
\* 35434 35533: gap of 100 bp  
\* 35534 36440: contig of 907 bp in length  
\* 36441 36540: gap of 100 bp  
\* 36541 37421: contig of 882 bp in length  
\* 37423 37522: gap of 100 bp  
\* 37523 38402: contig of 880 bp in length  
\* 38403 38502: gap of 100 bp  
\* 38503 39380: contig of 878 bp in length  
\* 39381 39480: gap of 100 bp  
\* 39481 40407: contig of 927 bp in length  
\* 40408 40507: gap of 100 bp  
\* 40508 41405: contig of 898 bp in length  
\* 41406 41505: gap of 100 bp  
\* 41506 42399: contig of 894 bp in length  
\* 42400 42499: gap of 100 bp  
\* 42500 43357: contig of 858 bp in length  
\* 43358 43457: gap of 100 bp  
\* 43458 44356: contig of 899 bp in length  
\* 44357 44456: gap of 100 bp  
\* 44457 45325: contig of 869 bp in length  
\* 45326 45425: gap of 100 bp  
\* 45426 46305: contig of 880 bp in length  
\* 46306 46405: gap of 100 bp  
\* 46406 47302: contig of 897 bp in length  
\* 47303 47402: gap of 100 bp  
\* 47403 48293: contig of 891 bp in length  
\* 48294 48393: gap of 100 bp  
\* 48394 49280: contig of 887 bp in length  
\* 49281 49380: gap of 100 bp  
\* 49381 50250: contig of 870 bp in length  
\* 50251 50350: gap of 100 bp  
\* 50351 51209: contig of 859 bp in length  
\* 51210 51309: gap of 100 bp  
\* 51310 52221: contig of 912 bp in length  
\* 52222 52321: gap of 100 bp  
\* 52322 53205: contig of 884 bp in length  
\* 53206 53305: gap of 100 bp  
\* 53306 54225: contig of 920 bp in length  
\* 54226 54325: gap of 100 bp  
\* 54326 55206: contig of 881 bp in length  
\* 55207 55306: gap of 100 bp  
\* 55307 56165: contig of 859 bp in length  
\* 56166 56265: gap of 100 bp  
\* 56266 57124: contig of 859 bp in length  
\* 57125 57224: gap of 100 bp  
\* 57225 58121: contig of 897 bp in length  
\* 58122 58221: gap of 100 bp  
\* 58222 59075: contig of 854 bp in length  
\* 59076 59175: gap of 100 bp  
\* 59176 60058: contig of 883 bp in length  
\* 60059 60158: gap of 100 bp  
\* 60159 61067: contig of 909 bp in length  
\* 61068 61167: gap of 100 bp  
\* 61168 62051: contig of 884 bp in length  
\* 62052 62151: gap of 100 bp  
\* 62152 63022: contig of 871 bp in length  
\* 63023 63122: gap of 100 bp  
\* 63123 64026: contig of 904 bp in length  
\* 64027 64126: gap of 100 bp  
\* 64127 65015: contig of 889 bp in length  
\* 65016 65115: gap of 100 bp  
\* 65116 65975: contig of 860 bp in length  
\* 65976 66075: gap of 100 bp  
\* 66076 66954: contig of 879 bp in length  
\* 66955 67054: gap of 100 bp  
\* 67055 67947: contig of 893 bp in length  
\* 67948 68047: gap of 100 bp  
\* 68048 68934: contig of 887 bp in length  
\* 68935 69034: gap of 100 bp  
\* 69035 69910: contig of 876 bp in length  
\* 69911 70010: gap of 100 bp

Query Match 4.7%; Score 152; DB 2; Length 90698;  
Best Local Similarity 99.2%; Pred. No. 5,1e-76;  
Matches 252; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 35 tactogggaggttgaggttgcaatgagccagatgcacattgcactccagcctgggca 94  
|||||  
Db 10177 TACTCGGAGGTGGAGGTGCAATGAGCCAGATGCACATTGCACCTCAGCCTGGCA 10236  
|||||  
QY 95 acaaaagtgaaactccatctcaattataaaaaaagaattttggtggtcgacttca 154  
|||||  
Db 10237 ACAAAGGTGAACCTCCATCTCAATTAATAAAAAAATAATGATTTGGTGGCGACTTCA 10296  
|||||  
QY 155 aatagtagagaagaagagagagagatgaggggtcaggagagatctaattactctca 214  
|||||  
Db 10297 AATAGTAGGAGAAGAAGAGAGAGAGATGGAGGTCACGGAGATCTAATTACTCTCTA 10356  
|||||  
QY 215 aatcatgctaggaagataaacaccttttaataaactctctgcttttataacatactc 274  
|||||  
Db 10357 AATCATGCTAGGAAGATAACACCTTTTATAAACACTCTCTGCTTTTATAACATCATTC 10416  
|||||  
QY 275 tgccaagagagctca 288  
|||||  
Db 10417 TGCCAAGGAGCTCA 10430  
|||||  
  
RESULT 30  
AX351034 AX351034 117 bp DNA linear PAT 06-FEB-2002  
LOCUS Sequence 6 from Patent WO0183746.  
ACCESSION AX351034  
VERSION AX351034.1 GI:18616390  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
  
REFERENCE 1 (sites)  
AUTHORS Rosier-Montus, M.F., Prades, C., Lemoine, C., Naudin, L., Deneffe, P.,  
Brewer, B., Duverger, N., Remaley, A. and Santamarina-Fojo, S.  
TITLE Regulatory nucleic acid sequences of the abcl gene  
JOURNAL Patent: WO 0183746-A 6 08-NOV-2001;  
Aventis Pharma S.A. (FR)  
FEATURES Location/Qualifiers  
source 1..117  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
  
BASE COUNT 20 a 30 c 34 g 33 t  
ORIGIN  
  
Query Match 3.6%; Score 117; DB 6; Length 117;  
Best Local Similarity 100.0%; Pred. No. 5,4e-56;  
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 3115 gtaagaggtctccagtgacttacttggtggcggtattgtttgttcgagccgaagagc 3174  
|||||  
Db 1 GTAAGAGGCTCTCCAGTGACTTACTTGGGCGTTATTGTTGTTGAGGCCAAGGAGGC 60  
|||||  
QY 3175 ttcgggaagtgcctcggttttcggggactttgatccggagccccacatccccaccatt 3231  
|||||  
Db 61 TTCGGGAAGTGCCTGGGTTTCGGGACTTTTGATCCGGAGCCCCACATCCCCACACTT 117  
|||||  
  
RESULT 31  
AX127764 AX127764 446 bp DNA linear PAT 15-MAY-2001  
LOCUS Sequence 3 from Patent WO0130848.  
ACCESSION AX127764  
VERSION AX127764.1 GI:14134411  
KEYWORDS synthetic construct.  
SOURCE synthetic construct  
ORGANISM artificial sequence.

REFERENCE 1 (bases 1 to 446)  
AUTHORS Deneffe,P., Rosier-Montus,M.F., Arnould-Reguigne,I., Prades,C.,  
Naudin,L., Lemoine,C., Duverger,N., Jaye,M., Searfoss,G.H.,  
Remaley,A., Brewer,H.B. and Dean,M.  
TITLE Nucleic acids of the human abcl gene and their therapeutic and  
diagnostic application  
JOURNAL Patent: WO 0130848-A 3 03-MAY-2001;  
Aventis Pharma S.A. (FR)  
FEATURES  
source Location/Qualifiers  
1. .446  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="Oligonucleotide primer"  
BASE COUNT 96 a 123 c 112 g 115 t  
ORIGIN  
Query Match 2.8%; Score 92; DB 6; Length 446;  
Best Local Similarity 100.0%; Pred. No. 2.1e-41;  
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3023 cttgtttttcccggtctgttttcccttccggaagcttgcgaagggttagga 3082  
|||||  
Db 1 CTTGTTTTTCCCGGTTCTGTTTTCTCCCTTCTCCGGAAGCTGTGTCAGGGGTAGGA 60  
QY 3083 gaaagagacgcaaacacaaagtggaaaacag 3114  
|||||  
Db 61 GAAAGAGACGCAACACAAAGTGGAAAAACAG 92  
RESULT 32  
AXI39751  
LOCUS AXI39751 446 bp DNA linear PAT 30-MAY-2001  
DEFINITION Sequence 3 from Patent EP1096012.  
ACCESSION AXI39751  
VERSION AXI39751.1 GI:14275333  
SOURCE synthetic construct.  
ORGANISM synthetic construct.  
REFERENCE 1 (bases 1 to 446)  
AUTHORS Deneffe,P., Rosier-Montus,M.F., Arnould-Reguigne,I., Prades,C.,  
Naudin,L., Lemoine,C., Duverger,N., Jaye,M., Searfoss Iii,G.H.,  
Remaley,A., Brewer,H.B. and Dean,M.  
TITLE Nucleic acids of the human abcl gene and their therapeutic and  
diagnostic application  
JOURNAL Patent: EP 1096012-A 3 02-MAY-2001;  
Aventis Pharma S.A. (FR)  
FEATURES  
source Location/Qualifiers  
1. .446  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="Oligonucleotide primer"  
BASE COUNT 96 a 123 c 112 g 115 t  
ORIGIN  
Query Match 2.8%; Score 92; DB 6; Length 446;  
Best Local Similarity 100.0%; Pred. No. 2.1e-41;  
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3023 cttgtttttcccggtctgttttcccttccggaagcttgcgaagggttagga 3082  
|||||  
Db 1 CTTGTTTTTCCCGGTTCTGTTTTCTCCCTTCTCCGGAAGCTGTGTCAGGGGTAGGA 60  
QY 3083 gaaagagacgcaaacacaaagtggaaaacag 3114  
|||||  
Db 61 GAAAGAGACGCAACACAAAGTGGAAAAACAG 92  
RESULT 33  
AXI27830  
LOCUS AXI27830 9741 bp DNA linear PAT 15-MAY-2001

DEFINITION Sequence 69 from Patent WO0130848.  
ACCESSION AXI27830  
VERSION AXI27830.1 GI:14134477  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 9741)  
AUTHORS Deneffe,P., Rosier-Montus,M.F., Arnould-Reguigne,I., Prades,C.,  
Naudin,L., Lemoine,C., Duverger,N., Jaye,M., Searfoss,G.H.,  
Remaley,A., Brewer,H.B. and Dean,M.  
TITLE Nucleic acids of the human abcl gene and their therapeutic and  
diagnostic application  
JOURNAL Patent: WO 0130848-A 69 03-MAY-2001;  
Aventis Pharma S.A. (FR)  
FEATURES  
source Location/Qualifiers  
1. .9741  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 2650 a 2180 c 2290 g 2620 t 1 others  
ORIGIN  
Query Match 2.8%; Score 92; DB 6; Length 9741;  
Best Local Similarity 100.0%; Pred. No. 2.9e-41;  
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3023 cttgtttttcccggtctgttttcccttccggaagcttgcgaagggttagga 3082  
|||||  
Db 1 CTTGTTTTTCCCGGTTCTGTTTTCTCCCTTCTCCGGAAGCTGTGTCAGGGGTAGGA 60  
QY 3083 gaaagagacgcaaacacaaagtggaaaacag 3114  
|||||  
Db 61 GAAAGAGACGCAACACAAAGTGGAAAAACAG 92  
RESULT 34  
AXI39817  
LOCUS AXI39817 9741 bp DNA linear PAT 30-MAY-2001  
DEFINITION Sequence 69 from Patent EP1096012.  
ACCESSION AXI39817  
VERSION AXI39817.1 GI:14275399  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 9741)  
AUTHORS Deneffe,P., Rosier-Montus,M.F., Arnould-Reguigne,I., Prades,C.,  
Naudin,L., Lemoine,C., Duverger,N., Jaye,M., Searfoss Iii,G.H.,  
Remaley,A., Brewer,H.B. and Dean,M.  
TITLE Nucleic acids of the human abcl gene and their therapeutic and  
diagnostic application  
JOURNAL Patent: EP 1096012-A 69 02-MAY-2001;  
Aventis Pharma S.A. (FR)  
FEATURES  
source Location/Qualifiers  
1. .9741  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 2650 a 2180 c 2290 g 2620 t 1 others  
ORIGIN  
Query Match 2.8%; Score 92; DB 6; Length 9741;  
Best Local Similarity 100.0%; Pred. No. 2.9e-41;  
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3023 cttgtttttcccggtctgttttcccttccggaagcttgcgaagggttagga 3082  
|||||  
Db 1 CTTGTTTTTCCCGGTTCTGTTTTCTCCCTTCTCCGGAAGCTGTGTCAGGGGTAGGA 60  
QY 3083 gaaagagacgcaaacacaaagtggaaaacag 3114

```

|||||
Db 61 GAAGAGACGCAACACAAAAGTGGAACACAG 92
|||||
RESULT 35
AX351038
LOCUS AX351038
DEFINITION Sequence 10 from Patent WO0183746.
ACCESSION AX351038
VERSION AX351038.1 GI:18616393
KEYWORDS
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (sites)
REFERENCE
AUTHORS Rosier-Montus, M.F., Prades, C., Lemoine, C., Naudin, L., Denefle, P.,
Brewer, B., Duverger, N., Remaley, A. and Santamarina-Fojo, S.
TITLE Regulatory nucleic acid sequences of the abcl1 gene
JOURNAL Patent: WO 0183746-A 10 08-NOV-2001;
Aventis Pharma S.A. (FR)
FEATURES
source
Location/Qualifiers
1. .9741
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2650 a 2180 c 2290 g 2620 t 1 others
ORIGIN
Query Match 2.8%; Score 92; DB 6; Length 9741;
Best Local Similarity 100.0%; Pred. No. 2.9e-41;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3023 cttgtttttcccggttctgttttcccttccttcggaaggtttcgaagggttagga 3082
|||||
Db 1 CTTGTTTTTCCCGGGTCTGTTTTCCTCCCTTCCTCCGGAAGGCTGTCAAGGGGTAGGA 60
|||||
Qy 3083 gaagagacgcaaacacaaaagtggaaaacag 3114
|||||
Db 61 GAAGAGACGCAACACAAAAGTGGAACACAG 92
|||||
RESULT 36
AC004139
LOCUS AC004139
DEFINITION Homo sapiens chromosome 17, clone HRP843B9, complete sequence.
ACCESSION AC004139
VERSION AC004139.1 GI:3513309
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 126295)
REFERENCE
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE Homo sapiens chromosome 17, clone HRP843B9
JOURNAL Unpublished
REFERENCE
AUTHORS Birren, B., Fasman, K., McKernan, K., Nusbaum, C., Richardson, P.,
Lander, E., Allen, N., Baker, J., Baldwin, J., Barna, N., Beckerly, R.,
Boutwell, C., Byrne, S., Cantu, C., Castle, A., Cerny, J., Cooke, P.,
Daly, M.J., Depayre, E., Devon, K., Dewar, K., Donelan, L., DuRette, B.,
Etamadi, S., Ferreira, P., Forrest, C., Funke, R., Gage, D., Gardyna, S.,
Gensheimer, S., Geraigery, K., Gilmartin, T., Gray, D., Hagos, B.,
Harris, K., Horton, L., Howland, J.C., Hui, L., Jacotot, L., Linton, L.,
MacKenzie, J., Marquis, N., McEwan, P., McGurk, A., Meldrim, J.,
Molla, M., Morris, W., Morrow, J., Nachman, A., Naylor, J., O'Connor, T.,
Paylin, B., Peterson, K., Ranganath, S., Riley, R., Roberts, D.,
Rollins, G., Rossello, R., Roy, A., Shyam, R., Soohoo, S.,
Stange-Thomann, N., Stilwell, J., Stone, C., Strickland, C., Sydney, K.,
Tang, L., Vassiliev, H., Vo, A., Wagner, A., Wheeler, J., Wu, Y.,
Ye, W.J., Zemtseva, I., Zhao, J. and Zody, M.
Direct Submission

```

```

JOURNAL Submitted (13-FEB-1998) Whitehead Institute/MIT Center for Genome
REFERENCE Research, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS 3 (bases 1 to 126295)
Birren, B., Fasman, K., Linton, L., Nusbaum, C., Lander, E., Allen, N.,
Anderson, M., Baker, J., Baldwin, J., Barna, N., Beckerly, R., Benn, J.,
Boutwell, C., Brown, A., Castle, A., Cerny, J., Collangelo, M.,
Collins, S., Collumore, A., Cooke, P., Corliss, D., Depayre, E.,
Devon, K., Dewar, K., Donelan, L., Ferreira, P., FitzHugh, W.,
Forrest, C., Funke, R., Gage, D., Gardyna, S., Geraigery, K., Grant, G.,
Hagos, B., Headford, A., Herena, L., Horton, L., Howland, J.C.,
Jacotot, L., Jones, C., Kann, L., Karatas, A., Lehoczy, J.,
Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,
Meldrim, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J.,
Nahf, R., Naylor, J., Niloff, M., O'Connor, T., O'Donnell, P.,
Paylin, B., Peterson, K., Riley, R., Roberts, D., Roy, A., Severy, P.,
Stange-Thomann, N., Stilwell, J., Stojanovic, N., Stone, C.,
Subramaniam, A., Tesfaye, S., Tichovolsky, N., Torruella-Miller, I.,
Vassiliev, H., Vo, A., Wagner, A., Wheeler, J., Wu, Y., Wyman, D.,
Ye, W.J., Zhao, J. and Zody, M.
Direct Submission
JOURNAL Submitted (02-SEP-1998) Whitehead Institute/MIT Center for Genome
REFERENCE Research, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS On Sep 2, 1998 this sequence version replaced gi:3451370.
COMMENT All repeats were identified using RepeatMasker: Smt, A.F.A. &
Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html.
FEATURES
source
Location/Qualifiers
1. .126295
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="HRPC843B9"
/repeat_region complement(2..193)
repeat_region /rpt_family="AluX"
repeat_region complement(235..410)
repeat_region /rpt_family="MER5A"
repeat_region complement(479..640)
repeat_region /rpt_family="MIR"
repeat_region complement(714..1012)
repeat_region /rpt_family="AluSc"
repeat_region 1140..1269
repeat_region /rpt_family="L2"
repeat_region complement(2503..2646)
repeat_region /rpt_family="MER5B"
repeat_region complement(2647..2954)
repeat_region /rpt_family="AluJb"
repeat_region complement(2955..2998)
repeat_region /rpt_family="MER5B"
repeat_region 3758..4544
repeat_region /rpt_family="L2"
repeat_region 4743..4873
repeat_region /rpt_family="FLAM_C"
repeat_region complement(5068..5364)
repeat_region /rpt_family="AluJb"
repeat_region complement(6403..6675)
repeat_region /rpt_family="AluSg"
repeat_region 6782..6950
repeat_region /rpt_family="MIR"
repeat_region complement(7379..7734)
repeat_region /rpt_family="Tigger1"
repeat_region complement(7735..7840)
repeat_region /rpt_family="(GGGA)n"
repeat_region complement(7841..7876)
repeat_region /rpt_family="purine-rich"
repeat_region complement(7883..7988)
repeat_region /rpt_family="Tigger1"
repeat_region complement(7989..8597)
repeat_region /rpt_family="Tigger1"
repeat_region complement(8604..8741)
repeat_region /rpt_family="FLAM_C"
repeat_region complement(8759..9456)

```

```
repeat_region /rpt_family="Tigger1"
complement(9467..9759)
/rpt_family="AluSx"
repeat_region complement(9820..10276)
/rpt_family="Tigger1"
repeat_region 10754..11095
/rpt_family="MER47A"
repeat_region complement(11106..11949)
/rpt_family="L2"
repeat_region 12033..12147
/rpt_family="L2"
repeat_region 12148..12235
/rpt_family="L1MC2"
repeat_region complement(13066..13222)
/rpt_family="L2"
repeat_region 13224..13251
/rpt_family="MIR"
repeat_region 13252..13544
/rpt_family="AluSg"
repeat_region 13545..13769
/rpt_family="MIR"
repeat_region 14208..14364
/rpt_family="MER5A"
repeat_region complement(14835..15195)
/rpt_family="L1MEC"
repeat_region complement(15353..15497)
/rpt_family="L2"
repeat_region complement(15662..15828)
/rpt_family="MLT1G"
repeat_region complement(16301..16607)
/rpt_family="AluY"
repeat_region 17500..17802
/rpt_family="AluSg"
repeat_region complement(18474..18774)
/rpt_family="AluSg"
repeat_region 18825..18902
/rpt_family="MIR"
repeat_region complement(19656..19771)
/rpt_family="MER20"
repeat_region 19834..20129
/rpt_family="AluSp"
repeat_region complement(20145..20168)
/rpt_family="(CAAAA)n"
repeat_region complement(20169..20451)
/rpt_family="AluSx"
repeat_region complement(20456..20765)
/rpt_family="AluSg"
repeat_region complement(21023..21343)
/rpt_family="L1M4"
repeat_region 21344..21630
/rpt_family="AluSx"
repeat_region complement(21631..21700)
/rpt_family="L1M4"
repeat_region 22291..22526
/rpt_family="MIR"
repeat_region 23606..23899
/rpt_family="AluJo"
repeat_region complement(24202..25601)
/rpt_family="L2"
repeat_region complement(25816..25978)
/rpt_family="L1MD"
repeat_region 26161..26465
/rpt_family="AluSx"
repeat_region complement(26630..26725)
/rpt_family="MIR"
repeat_region complement(26902..26958)
/rpt_family="(CA)n"
repeat_region complement(26959..27179)
/rpt_family="Aluub"
repeat_region complement(27189..27489)
/rpt_family="AluSg"
repeat_region 27712..28009
/rpt_family="AluSg"
```

```
complement(28773..29080)
/rpt_family="AluSx"
repeat_region 29132..29580
/rpt_family="L1MCS"
repeat_region 29962..30097
/rpt_family="L2"
repeat_region 30275..30566
/rpt_family="AluSx"
repeat_region complement(31041..31353)
/rpt_family="AluSg"
repeat_region complement(32362..32571)
/rpt_family="AluSg/x"
repeat_region complement(33470..33781)
/rpt_family="AluYb8"
repeat_region complement(35504..35774)
/rpt_family="L1M3A"
repeat_region 35830..35906
/rpt_family="AluJo"
repeat_region 35907..36194
/rpt_family="AluY"
repeat_region 36204..36516
/rpt_family="AluSg"
repeat_region 36517..36682
/rpt_family="AluJo"
repeat_region complement(36689..37629)
```

Query Match 2.0%; Score 65; DB 9; Length 126295;  
Best Local Similarity 100.0%; Pred. No. 1.8e-25;  
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1017 gatctcgactgaactgtacctgtcctcccggttaagcgattctctgcctcagctc 1076  
|||||  
Db 31115 GATCTCGACTGACTGTAACTCTGCTCCCGGTTCAAGCGATTCTCTGCTCAGCTC 31174  
|||||

QY 1077 ctgag 1081  
|||||

Db 31175 CTGAG 31179

RESULT 37  
AC021850/c

LOCUS AC021850 206966 bp DNA linear PRI 09-MAY-2001  
DEFINITION Homo sapiens BAC clone RP11-389E17 from 4, complete sequence.  
ACCESSION AC021850  
VERSION AC021850.8 GI:12863232  
KEYWORDS HTG.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
REFERENCE  
2 (bases 1 to 206966)  
Kang, K., Maupin, R. and Ureta, M.  
The sequence of Homo sapiens BAC clone RP11-389E17  
Unpublished  
3 (bases 1 to 206966)  
Waterston, R.H.  
Direct Submission  
Submitted (20-JAN-2000) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
4 (bases 1 to 206966)  
Waterston, R.H.  
Direct Submission  
Submitted (16-FEB-2001) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
5 (bases 1 to 206966)  
Waterston, R.

TITLE Direct Submission  
 JOURNAL Submitted (09-MAY-2001) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 COMMENT On Feb 16, 2001 this sequence version replaced gi:111761497.  
 ----- Genome Center  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: <http://genome.wustl.edu/gsc>  
 Contact: saplens@wustl.wustl.edu  
 ----- Summary Statistics  
 Center project name: H\_NH0389E17  
 -----

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

#### MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

#### SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tateno, M., Catanese, J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)  
 VECTOR: pBACe3.6

#### NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-62B4. Actual start of this clone is at base position 1 of RP11-389E17; actual end is at base position 206966 of RP11-389E17.

#### FEATURES

source	Location/Qualifiers
	1. .206966 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="4" /map="4"
repeat_region	/clone="RP11-389E17"
repeat_region	/clone_lib="RPCI-11"
repeat_region	3. .256 /rpt_family="L1"
repeat_region	389. .509 /rpt_family="L1"
repeat_region	523. .1023 /rpt_family="L1"
repeat_region	1088. .1319 /rpt_family="L1"
repeat_region	1779. .1908 /rpt_family="MIR"
repeat_region	1930. .1950 /rpt_family="AT_rich"
repeat_region	2531. .2753 /rpt_family="Alu"
repeat_region	2772. .2792 /rpt_family="AT_rich"
repeat_region	2987. .3022 /rpt_family="AT_rich"
repeat_region	3223. .3260 /rpt_family="L1"
	/rpt_family="AT_rich"
	3386. .3523 /rpt_family="MIR"
repeat_region	3775. .4559 /rpt_family="ERV1"
repeat_region	4595. .4763 /rpt_family="ERV1"
repeat_region	5163. .5193 /rpt_family="(CA)n"
repeat_region	5276. .5297 /rpt_family="AT_rich"
repeat_region	7864. .8032 /rpt_family="MIR"
repeat_region	8397. .8660 /rpt_family="ERV1"
repeat_region	8877. .8991 /rpt_family="GA-rich"
repeat_region	9049. .9235 /rpt_family="L2"
repeat_region	9284. .9364 /rpt_family="L1"
repeat_region	9388. .9734 /rpt_family="L1"
repeat_region	11490. .11628 /rpt_family="L2"
repeat_region	12620. .12715 /rpt_family="MER2_type"
repeat_region	12726. .12923 /rpt_family="MaLR"
repeat_region	12950. .13462 /rpt_family="MER2_type"
repeat_region	17273. .17508 /rpt_family="Alu"
repeat_region	18160. .18215 /rpt_family="L2"
repeat_region	18689. .18810 /rpt_family="MIR"
repeat_region	19217. .19324 /rpt_family="L1"
repeat_region	20012. .20292 /rpt_family="L1"
repeat_region	20476. .20830 /rpt_family="MaLR"
repeat_region	23028. .23085 /rpt_family="AT_rich"
repeat_region	23483. .23612 /rpt_family="Achoho"
repeat_region	23878. .23976 /rpt_family="AT_rich"
repeat_region	24263. .24404 /rpt_family="MIR"
repeat_region	24706. .24903 /rpt_family="MIR"
repeat_region	26677. .27207 /rpt_family="MER1_type"
repeat_region	27597. .27619 /rpt_family="AT_rich"
repeat_region	28709. .28781 /rpt_family="AT_rich"
repeat_region	29420. .29635 /rpt_family="Alu"
repeat_region	29865. .29898 /rpt_family="tRNA-Cys-TGC"
repeat_region	29979. .30116 /rpt_family="MIR"
repeat_region	30773. .31036 /rpt_family="L2"
repeat_region	31039. .31722 /rpt_family="L2"
repeat_region	32160. .32333 /rpt_family="CRI"
repeat_region	33594. .34090 /rpt_family="L1"

```

repeat_region 34248..34275
/rpt_family="AT_rich"
repeat_region 34635..34825
/rpt_family="L1"
repeat_region 35013..35078
/rpt_family="AT_rich"
repeat_region 35645..35691
/rpt_family="(CA)n"
repeat_region 36784..37028
/rpt_family="L1"
repeat_region 37100..37171
/rpt_family="L1"
repeat_region 37381..37431
/rpt_family="AT_rich"
repeat_region 37582..38445
/rpt_family="ERVL"
repeat_region 39128..39148
/rpt_family="AT_rich"
repeat_region 40735..40755
/rpt_family="AT_rich"
repeat_region 40826..41030
/rpt_family="L1"

Query Match
Best Local Similarity 100.0%; Pred. No. 4e-23; Length 206966;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1017 gatcgactcactgaacctgctccgggttaacgagattctcctgcctcagcctc 1076
|||||
Db 45729 GATCTGACTCACTGTAACCTGCTCCGGGTTCACGGATTCCCTCAGCCTC 45670

QY 1077 c 1077
Db 45669 c 45669

RESULT 38
AF287263
LOCUS AF287263 278572 bp DNA linear ROD 23-APR-2001
DEFINITION Mus musculus ATP-binding cassette 1, sub-family A, member 1 (Abca1)
ACCESSION AF287263
VERSION AF287263.1 GI:11611824
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE Qiu,Y., Cavellier,L., Chiu,S., Yang,X., Rubin,E., and Cheng,J.F.
1 (bases 1 to 278572)
TITLE Human and mouse abca1 comparative sequencing and transgenesis
studies revealing novel regulatory sequences
JOURNAL Genomics 73 (1), 66-76 (2001)
MEDLINE 21251004
AUTHORS Qiu,Y., Cavellier,L., Chiu,S., Rubin,E. and Cheng,J.-F.
REFERENCE 2 (bases 1 to 278572)
TITLE Direct Submission
JOURNAL Submitted (14-JUL-2000) Genome Science Department, Lawrence
Berkeley National Laboratory, 1 Cyclotron Rd, MS 84-171, Berkeley,
CA 94720, USA
FEATURES
source
1. 278572
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="BAC 129K10; 142M4; 197F14"
join(87184..87402,103551..103699,115064..115187,
119217..119358,120364..120482,133479..133600,
135582..135758,147083..147175,149791..150031,
152367..152506,152845..152961,156570..156767,
157459..157664,158200..158376,158821..159043,
160617..160838,161790..161994,163076..163189,
164646..164817,165738..165869,167171..167313,
167492..167629,168320..168540,169790..169862,
170735..170937,172636..172684,172855..172968,
173925..174073,175605..175729,176761..176859,
180181..180370,181179..181273,182395..182427,
183398..183503,191761..191829,192306..192475,
193631..193808,196619..196734,196845..196989,
198526..198649,200922..201051,201281..201401,
202114..202176,202890..202996,204955..205096,
205427..205558,206739..206842,207361..207453,
208135..208378,209638..209778)
/gene="Abca1"
/product="ATP-binding cassette 1, sub-family A, member 1"
87184..210783
/gene="Abca1"
join(119237..119358,120364..120482,133479..133600,
135582..135758,147083..147175,149791..150031,
152367..152506,152845..152961,156570..156767,
157459..157664,158200..158376,158821..159043,
160617..160838,161790..161994,163076..163189,
164646..164817,165738..165869,167171..167313,
167492..167629,168320..168540,169790..169862,
170735..170937,172636..172684,172855..172968,
173925..174073,175605..175729,176761..176859,
180181..180370,181179..181273,182395..182427,
183398..183503,191761..191829,192306..192475,
193631..193808,196619..196734,196845..196989,
198526..198649,200922..201051,201281..201401,
202114..202176,202890..202996,204955..205096,
205427..205558,206739..206842,207361..207453,
208135..208378,209638..209778)
/gene="Abca1"
/codon_start=1
/product="ATP-binding cassette 1, sub-family A, member 1"
/protein_id="AAG39073.1"
/db_xref="GI:11611825"
/translaton="MPSAGFLPWQGIICNANNPCRYPTTPGAPGVGVGNFKSIYSR
LFSDAQLLYSQDTSIKDMHKLRLQIKHPNSNLKQDFLVNFTSGFLQHLN
KLSPRTVDSLLQANVGLQYQLHSLACNGSKLEIIQIOLGAEVSAALGCLPRK
KLDAARVLYRNMIDILKPVVTKLNSHLTPHLEATVLLDLSGLAQELPSTKSK
SDMOEYMETLVNSSSSQIYQAVSRIVCGHGGGLKIKSLNVEONNYKALRGG
NTEEDVDTFYDNTTPYCNLMKNLESSPLSRIWKALKPLLVGLKILTPDPATRO
VMAEVNKTFOELAVFHLEGMWELSPQIWMFMSQEMDLVRLTLDSDGNDWFQSK
LDGLDTAQDIMAFLAKNEDVQSPNGSVYTRAFNETNQAIQTISRMECINLVNKL
EPIPTVRLINKSMELLDERKWAGVIFTGTPDSVELPHVYKIRKMDINDVNTNK
IKDGYWDPGRADPFEDMYVGGFAYLDQVBOAIIRVLTGSEKTYVVOQMPYPC
YVDIDFLRVMSRNPFLMTLAWISYAVI IKSIYKEARLKETMRIMGLDNLWFS
WFVSSLIPLLVASGLLVILKLNLLPYSDPSVVFVLSVFMVTLQCFELISTFSR
ANLAAGCGGIYFTLYLPVLCVAMQDYGVFSIKFASLSVPAVFGCEYFALFEEQ
GIGQWNLFPESPEDEGNLTATVASMFLDFTLYGVMTWYIEAVFPQGYGPRPYF
PCTKSYWFEIDEIKSHPGSQKSGVEICMEEPHTHLRGVSIQNLVYRDMKVVAV
DGLALNFEGQITSLFHNGAGKTTMSITGLPTSGTAYILGDKIRSEMSIRQN
LGVCPOHNVLFDMTLVEEHLWFYARLKGSEKHVKAEMQALDGLPPLSKLSTSQ
LSGSMOKLSVALAFVGGSKVILDEPTAGVDPSYRSGIWEKLLYRQGTIILSTHH
MDEADILGRIALISHGKCCVSSSLFNQLGTGYLILVKKDVSSSSCKNSST
VSLCKEDSVSSQSSDAGSDHSDTLTIDVSAISNLIRKHVSEARVEDIGHETY
VLPYEAKEGAFVELFHEIDRSLDGLISYSETLEEFILKVAEESGVDAETSDG
TLPARNRRAFGDKQLPFTEDDADPNDSIDIPESRETDLLSGMDKSGVQLKGW
KLTOQOQVALLKRLIARRSRKGFPAIIVPAVFCIALVFLIIVPGRKVPSELQ
PMWNYQYTFVSDNAPEDMGTQELLNALTQDGFGRCMGPNIPDPCLAGEDWTI
SPVQSIIVLDLQNGNTMKNPACQSSDKIKMLPVCPGPGAGLPPQPKQKTDI
LQNLTRISDLYLVKTYVQLIAKSLKNKIWNFEYRGYSGLVSGVSNAPPSHEVNA
IKOMKLLKLTADRSFSLGRFMAGLDTKNNVYFNKNGHWHAISFLNINNAIIL
RANLQGENPQYQITAFNHPNLTKQLSEVALMTYSDVLVSIQVIFAMSGVPAS
VFNLIQBRVSKAKHLOFISGVPIVYLVNFVDMCNVYVPATLVIIIFICQOKSYV
SSTNLPVALIILLLXGWSITPLMYPASFVKIPSTAVVLTIVNLFINGSVATPVL
ELFTNNKLDINDILKSVFLIPFHCCLGRLIDMKVKNQAMADALERGNRRVPSLSW
DLVGRNLFAWEGVVFLLTILQIRFIRPVRKAKLPPLNDEGVRERQRILTD
GGGNDILTEIKLTIKYRKRKPADVRCIGIPPGCEGLLVNGAGKSTTFKMTGD
TPVTRGDALFNKLSILNIEHVQNNQYCPQFADITELLTGREHVEFFALLRGVPEKE
VGKGEWAIRKGLVYKGYKAYNSGKNRKLSTAMALIGGPVPLVDEPTGMGPKA
RRFLWVRIAGSNPDLPKPOVEFFGLAPPGSVLKEKHNMQLQYLPSSLSLARIFSLLS
GTYIVRIAGSNPDLPKPOVEFFGLAPPGSVLKEKHNMQLQYLPSSLSLARIFSLLS
QSKKRHLIEDIYSVSTQLDQVFNFAKQSDDDHLKDLSLHKNQTVVDVAVSLQDQ

```



```

BASE COUNT      76801 a 58762 c 61256 g 81498 t 255 others
ORIGIN

Query Match      1.9% Score 61; DB 10; Length 278572;
Best Local Similarity 100.08; Pred. No. 4-le-23;
Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2744 ctccacgtcttctgctgactgaactacataaagagccgggaagggcgccggg 2803
|||||
Db 87033 CTCCAGCTCTTCTGCTGAGTGACTGAACATACATAACAGAGCGCGGAGGGCGCGG 87092

QY 2804 g 2804
Db 87093 g 87093

RESULT 39
AC107969/c
LOCUS
DEFINITION      Homo sapiens chromosome 11 clone CTD-2028015 map 11, LOW-PASS
SEQUENCE SAMPLING.
ACCESSION      AC107969
VERSION
KEYWORDS
SOURCE
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 57662)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 11, clone CTD-2028015
Unpublished
2 (bases 1 to 57662)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepe,Y., Colangelo,M., Collins,S., Collamore,A., Cook,A.,
Cook,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferrel,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C.,
Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., Meldrum,J., Meneus,L., Minova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembeke,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L24518
Center Clone name: 2028_O_15
-----
* NOTE: This record contains 73 individual

```

```

* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
*
* 1 717: contig of 717 bp in length
* 718 817: gap of 100 bp
* 818 1525: contig of 708 bp in length
* 1526 1625: gap of 100 bp
* 1626 2265: contig of 640 bp in length
* 2266 2365: gap of 100 bp
* 2366 3050: contig of 685 bp in length
* 3051 3150: gap of 100 bp
* 3151 3845: contig of 695 bp in length
* 3846 3945: gap of 100 bp
* 3946 4577: contig of 632 bp in length
* 4578 4677: gap of 100 bp
* 4678 5369: contig of 692 bp in length
* 5370 5469: gap of 100 bp
* 5470 6151: contig of 682 bp in length
* 6152 6251: gap of 100 bp
* 6252 6955: contig of 704 bp in length
* 6956 7055: gap of 100 bp
* 7056 7804: contig of 749 bp in length
* 7805 7904: gap of 100 bp
* 7905 8603: contig of 699 bp in length
* 8604 8703: gap of 100 bp
* 8704 9392: contig of 689 bp in length
* 9393 9492: gap of 100 bp
* 9493 10155: contig of 663 bp in length
* 10156 10255: gap of 100 bp
* 10256 10948: contig of 693 bp in length
* 10949 11048: gap of 100 bp
* 11049 11603: contig of 555 bp in length
* 11604 11703: gap of 100 bp
* 11704 12386: contig of 683 bp in length
* 12387 12486: gap of 100 bp
* 12487 13187: contig of 701 bp in length
* 13188 13287: gap of 100 bp
* 13288 13988: contig of 701 bp in length
* 13989 14088: gap of 100 bp
* 14089 14795: contig of 707 bp in length
* 14796 14895: gap of 100 bp
* 14896 15532: contig of 637 bp in length
* 15533 15632: gap of 100 bp
* 15633 16347: contig of 715 bp in length
* 16348 16447: gap of 100 bp
* 16448 17155: contig of 708 bp in length
* 17156 17255: gap of 100 bp
* 17256 17961: contig of 706 bp in length
* 17962 18061: gap of 100 bp
* 18062 18723: contig of 662 bp in length
* 18724 18823: gap of 100 bp
* 18824 19476: contig of 653 bp in length
* 19477 19576: gap of 100 bp
* 19577 20274: contig of 698 bp in length
* 20275 20374: gap of 100 bp
* 20375 21074: contig of 700 bp in length
* 21075 21174: gap of 100 bp
* 21175 21867: contig of 693 bp in length
* 21868 21967: gap of 100 bp
* 21968 22680: contig of 713 bp in length
* 22681 22780: gap of 100 bp
* 22781 23495: contig of 715 bp in length
* 23496 23595: gap of 100 bp
* 23596 24227: contig of 632 bp in length
* 24228 24327: gap of 100 bp
* 24328 25032: contig of 705 bp in length

```

\* 25033 25132: gap of 100 bp  
\* 25133 25838: contig of 706 bp in length  
\* 25839 25938: gap of 100 bp  
\* 25939 26651: contig of 713 bp in length  
\* 26652 26751: gap of 100 bp  
\* 26752 27431: contig of 680 bp in length  
\* 27432 27531: gap of 100 bp  
\* 27532 28232: contig of 701 bp in length  
\* 28233 28332: gap of 100 bp  
\* 28333 29032: contig of 700 bp in length  
\* 29033 29132: gap of 100 bp  
\* 29133 29838: contig of 706 bp in length  
\* 29839 29938: gap of 100 bp  
\* 29939 30653: contig of 715 bp in length  
\* 30654 30753: gap of 100 bp  
\* 30754 31466: contig of 713 bp in length  
\* 31467 31566: gap of 100 bp  
\* 31567 32272: contig of 706 bp in length  
\* 32273 32372: gap of 100 bp  
\* 32373 33009: contig of 637 bp in length  
\* 33010 33109: gap of 100 bp  
\* 33110 33827: contig of 718 bp in length  
\* 33828 33927: gap of 100 bp  
\* 33928 34646: contig of 719 bp in length  
\* 34647 34746: gap of 100 bp  
\* 34747 35451: contig of 705 bp in length  
\* 35452 35551: gap of 100 bp  
\* 35552 36244: contig of 693 bp in length  
\* 36245 36344: gap of 100 bp  
\* 36345 36993: contig of 649 bp in length  
\* 36994 37093: gap of 100 bp  
\* 37094 37777: contig of 684 bp in length  
\* 37778 37877: gap of 100 bp  
\* 37878 38587: contig of 710 bp in length  
\* 38588 38687: gap of 100 bp  
\* 38688 39375: contig of 688 bp in length  
\* 39376 39475: gap of 100 bp  
\* 39476 40179: contig of 704 bp in length  
\* 40180 40279: gap of 100 bp  
\* 40280 40986: contig of 707 bp in length  
\* 40987 41086: gap of 100 bp  
\* 41087 41749: contig of 663 bp in length  
\* 41750 41849: gap of 100 bp  
\* 41850 42551: contig of 702 bp in length  
\* 42552 42651: gap of 100 bp  
\* 42652 43363: contig of 712 bp in length  
\* 43364 43463: gap of 100 bp  
\* 43464 44144: contig of 681 bp in length  
\* 44145 44244: gap of 100 bp  
\* 44245 44939: contig of 695 bp in length  
\* 44940 45039: gap of 100 bp  
\* 45040 45740: contig of 701 bp in length  
\* 45741 45840: gap of 100 bp  
\* 45841 46513: contig of 673 bp in length  
\* 46514 46613: gap of 100 bp  
\* 46614 47307: contig of 694 bp in length  
\* 47308 47407: gap of 100 bp  
\* 47408 48097: contig of 690 bp in length  
\* 48098 48197: gap of 100 bp  
\* 48198 48884: contig of 687 bp in length  
\* 48885 48984: gap of 100 bp  
\* 48985 49700: contig of 716 bp in length  
\* 49701 49800: gap of 100 bp  
\* 49801 50484: contig of 684 bp in length  
\* 50485 50584: gap of 100 bp  
\* 50585 51288: contig of 704 bp in length  
\* 51289 51388: gap of 100 bp  
\* 51389 52107: contig of 719 bp in length  
\* 52108 52207: gap of 100 bp  
\* 52208 52909: contig of 702 bp in length  
\* 52910 53009: gap of 100 bp  
\* 53010 53665: contig of 656 bp in length  
\* 53666 53765: gap of 100 bp

\* 53766 54444: contig of 679 bp in length

Query Match 1.8%; Score 58; DB 2; Length 57662;  
Best Local Similarity 100.0%; Pred. No. 1.9e-21;  
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1187 ctgccacattgctcccaagtctgtgattacagcatgagccactgcgccagc 1244  
|||||  
DB 39017 CTGCCACCTTGGCTCCCAAAGTCTGGATTACAGCATGAGCCACTGCCCCAGC 38960  
|||||

RESULT 40  
HS931E15/c

LOCUS  
DEFINITION  
Human DNA sequence from clone 931E15 on chromosome Xq25. Contains  
STSs, GSSs and genomic marker DXS8098, complete sequence.

ACCESSION  
AL023575

VERSION  
AL023575.1 GI:3618163

KEYWORDS  
HTG; DXS8098.

SOURCE  
human.

ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 81874)

REFERENCE  
AUTHORS  
Pavitt,R.

JOURNAL  
Direct Submission  
Submitted (24-SEP-1998) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
requests: clonerequest@sanger.ac.uk

COMMENT  
On Sep 18, 1998 this sequence version replaced gi:3550203.  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
This sequence has been finished according to sequence map criteria  
as follows. An attempt is made to resolve all sequencing problems,  
such as compressions and repeats, but not necessarily within known,  
annotated human repeat sequence elements (e.g. Alu). Where the  
sequence is ambiguous, there is an annotation using the 'unsure'  
feature key.  
This sequence was generated from part of bacterial clone contigs of  
human chromosome X, constructed by the Sanger Centre Chromosome X  
Mapping Group. Further information can be found at  
<http://www.sanger.ac.uk/HGP/ChrX>  
931E15 is from the library RPC15 constructed at the Roswell Park  
Cancer Institute by the group of Pieter de Jong. For further  
details see <http://bacpac.med.buffalo.edu/VECTOR:pcypac2>  
IMPORTANT: This sequence is not the entire insert of clone 931E15.  
It may be shorter because we only sequence overlapping sections  
once, or longer because we arrange for a small overlap between  
neighbouring submissions.  
The true left end of 424J12 (282207) is at 36643 in this sequence.  
The true right end of 506G2 (282213) is at 37431.

FEATURES  
source  
1. .81874  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="X"  
/map="q25"  
/clone="RP5-931E15"  
/cclone\_lib="RPC1-5"  
1. .201  
/note="AluX repeat: matches 94. .294 of consensus"  
complement(228. .531)  
/note="AluX repeat: matches 292. .1 of consensus"  
569. .687  
/note="AluJo repeat: matches 11. .127 of consensus"  
702. .1002  
/note="AluSp repeat: matches 1. .302 of consensus"  
1010. .1178  
/note="FRAM repeat: matches 5. .166 of consensus"

misc\_feature complement(1365..1664)  
/note="match: STS G05254"  
repeat\_region 2391..2691  
/note="AluSx repeat: matches 1..300 of consensus"  
repeat\_region 2767..3067  
/note="AluJb repeat: matches 1..302 of consensus"  
repeat\_region complement(4220..4537)  
/note="AluJb repeat: matches 300..1 of consensus"  
repeat\_region 4701..5004  
/note="AluJb repeat: matches 2..302 of consensus"  
repeat\_region 5170..5497  
/note="L1 repeat: matches 3627..3966 of consensus"  
repeat\_region 5498..5643  
/note="AluSx repeat: matches 1..146 of consensus"  
repeat\_region 5644..5948  
/note="AluY repeat: matches 1..301 of consensus"  
repeat\_region 5950..6115  
/note="AluSg repeat: matches 133..290 of consensus"  
repeat\_region 6156..6393  
/note="L1 repeat: matches 3985..4239 of consensus"  
repeat\_region 6629..6979  
/note="L1 repeat: matches 4377..4731 of consensus"  
repeat\_region 6978..7073  
/note="MER2 repeat: matches 1..95 of consensus"  
repeat\_region complement(7078..7139)  
/note="MER2 repeat: matches 345..284 of consensus"  
repeat\_region 7238..7675  
/note="L1 repeat: matches 4728..5161 of consensus"  
repeat\_region 7678..7969  
/note="AluSx repeat: matches 1..293 of consensus"  
repeat\_region 7970..8091  
/note="L1 repeat: matches 5150..5270 of consensus"  
repeat\_region 8094..8396  
/note="AluY repeat: matches 1..300 of consensus"  
repeat\_region 8400..8615  
/note="L1 repeat: matches 47..276 of consensus"  
repeat\_region 8616..8733  
/note="AluJb repeat: matches 2..121 of consensus"  
repeat\_region 8743..8816  
/note="AluY repeat: matches 228..301 of consensus"  
repeat\_region 8856..9146  
/note="AluJb repeat: matches 2..300 of consensus"  
repeat\_region 9152..9362  
/note="AluJb repeat: matches 87..299 of consensus"  
repeat\_region 9365..9667  
/note="AluJb repeat: matches 1..302 of consensus"  
repeat\_region 9668..10289  
/note="L1 repeat: matches 297..919 of consensus"  
repeat\_region complement(10371..10670)  
/note="AluSx repeat: matches 302..3 of consensus"  
repeat\_region 10710..11004  
/note="AluSx repeat: matches 7..301 of consensus"  
repeat\_region 11299..11939  
/note="L1 repeat: matches 271..910 of consensus"  
repeat\_region 11983..12087  
/note="L1 repeat: matches 6..109 of consensus"  
repeat\_region 12100..12401  
/note="AluSx repeat: matches 4..302 of consensus"  
repeat\_region 12439..12940  
/note="L1 repeat: matches 166..685 of consensus"  
misc\_feature complement(13266..13588)  
/note="match: STS L40936"  
repeat\_region 13604..13904  
/note="AluSx repeat: matches 2..302 of consensus"  
repeat\_region 13958..14050  
/note="AluSx repeat: matches 1..93 of consensus"  
repeat\_region 14055..14323  
/note="AluJb repeat: matches 60..301 of consensus"  
repeat\_region 14514..14836  
/note="AluJb repeat: matches 13..302 of consensus"  
repeat\_region 15126..15428  
/note="AluSx repeat: matches 1..302 of consensus"  
repeat\_region 15626..15730

repeat\_region /note="AluJb repeat: matches 196..300 of consensus"  
16118..16420  
/note="AluJb repeat: matches 1..290 of consensus"  
complement(16828..16935)  
/note="FLAM\_A repeat: matches 117..6 of consensus"  
complement(17160..17224)  
/note="MIR repeat: matches 149..84 of consensus"  
18120..18226  
/note="match: STS L24586"  
complement(18124..18827)  
/note="match: GSSs B16853 B14990"  
complement(19378..19676)  
/note="AluSp repeat: matches 303..2 of consensus"  
19901..20160  
/note="AluY repeat: matches 39..298 of consensus"  
complement(20716..20882)  
/note="MIR repeat: matches 260..69 of consensus"  
complement(20824..21385)  
/note="match: GSSs B94184 A0010265"  
21550..23628  
/note="putative CpG island"  
complement(22542..22581)  
/note="MIR2 repeat: matches 134..95 of consensus"  
complement(23094..24185)  
/note="SVA repeat: matches 1101..1 of consensus"  
24186..24231  
/note="23 copies 2 mer aa 80% conserved"  
24520..24884  
/note="L1 repeat: matches 517..890 of consensus"  
complement(25504..25687)  
/note="MER20 repeat: matches 155..1 of consensus"  
26421..26484  
/note="32 copies 2 mer tt 83% conserved"  
complement(26486..26661)  
/note="AluJb repeat: matches 195..5 of consensus"  
26867..27164  
/note="AluY repeat: matches 3..300 of consensus"  
27165..27462  
/note="AluSp repeat: matches 2..302 of consensus"  
28080..28392  
/note="AluSx repeat: matches 1..298 of consensus"  
complement(28817..29089)  
/note="AluSx repeat: matches 298..2 of consensus"  
29122..29252  
/note="L1 repeat: matches 3151..3280 of consensus"  
complement(29957..30200)  
/note="MLT2\_internal repeat: matches 5375..5124 of consensus"  
complement(30513..30706)  
/note="MLT2\_internal repeat: matches 4754..4572 of consensus"  
30728..31025  
/note="AluSx repeat: matches 1..299 of consensus"  
31059..31356  
/note="AluSg repeat: matches 1..296 of consensus"  
31359..31505  
/note="MER4B repeat: matches 11..148 of consensus"  
32041..32319  
/note="AluSx repeat: matches 1..287 of consensus"  
complement(32381..33599)  
/note="MLT2\_internal repeat: matches 4572..3316 of consensus"  
33603..33906  
/note="AluJb repeat: matches 1..302 of consensus"  
34020..34321  
/note="AluJb repeat: matches 1..301 of consensus"  
complement(34388..34933)

Query Match 1.8%; Score 58; DB 9; Length 81874;  
Best Local Similarity 100.0%; Pred. No. 2e-21;  
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1187 ctgccaccttgccctcccaagtgtgtgattacaggaatgagccactgcgccagc 1244

Db 39890 CTGCCACCTGGCGCTCCCAAAGTGC7GGGATTACAGGCATGAGCCACTGCGCCACG 398933  
 |||  
 RESULT 41  
 HSDJ247C2/c  
 LOCUS HSDJ247C2 98360 bp DNA linear PRI 03-AUG-2001  
 DEFINITION Human DNA sequence from clone RPI-247C2 on chromosome 11p13  
 Contains STSs and GSSs, complete sequence.  
 ACCESSION AL049713  
 VERSION AL049713.20 GI:10334639  
 KEYWORDS HTG.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 98360)  
 AUTHORS Wallis,J.  
 TITLE Direct Submission  
 JOURNAL Submitted (02-AUG-2001) Sanger Centre, Hinxton, Cambridgeshire,  
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone  
 requests: clonerequest@sanger.ac.uk  
 COMMENT On Sep 27, 2000 this sequence version replaced gi:6006528.  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.  
 The following abbreviations are used to associate primary accession  
 numbers given in the feature table with their source databases:  
 Em.; EMBL; SW.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information  
 on the WORMPEP database can be found at  
 http://www.sanger.ac.uk/Projects/C\_elegans/wormpep/IMPORTANT: This  
 sequence is not the entire insert of clone RPI-247C2 It may be  
 shorter because we sequence overlapping sections only once, except  
 for a 100 base overlap.  
 The true left end of clone RPI-247C2 is at 1 in this sequence. The  
 true left end of clone RP5-879J18 is at 98261 in this sequence. The  
 true right end of clone RPI-85W6 is at 4651 in this sequence. This  
 sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >=  
 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest. RPI-247C2 is from the  
 library RPCI-1 constructed by the group of Pieter de Jong. For  
 further details see  
 http://www.chori.org/bacpac/home.htm  
 VECTOR: pCVPAC2.  
 FEATURES  
 source  
 1..98360  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="11"  
 /map="p13"  
 /clone="RPI-247C2"  
 /clone\_lib="RPCI-1"  
 complement(627..983)  
 /note="match: STS: Em:HS198YB10"  
 misc\_feature  
 804..829  
 /note="13 copies 2 mer gt 100% conserved"  
 repeat\_region  
 3773..3810  
 /note="L1PAl1 repeat: matches 2154..2191 of consensus"  
 repeat\_region  
 3956..3992  
 /note="L1PAl1 repeat: matches 2180..2215 of consensus"  
 repeat\_region  
 4265..5286  
 /note="L1PAl1 repeat: matches 2215..3216 of consensus"  
 repeat\_region  
 5386..8453  
 /note="L1PAl1 repeat: matches 3216..6162 of consensus"  
 repeat\_region  
 10697..10906  
 /note="L1MEC repeat: matches 2142..1979 of consensus"

```

repeat_region 63917..64021
/note="LIMD repeat: matches 984..1089 of consensus"
/note="complement(64074..64543)
/note="match: GSS: Em:AQ885350"
/note="complement(66096..66577)
/note="match: GSS: Em:AQ611651"
69990..70342
/note="match: SPS: Em:G21603"
complement(76094..76635)
/note="match: GSS: Em:AQ277578"
complement(76184..76636)
/note="match: GSS: Em:AQ812553"
complement(76460..76628)
/note="match: GSS: Em:AQ140871"
complement(88581..89052)
/note="match: GSS: Em:AQ201907"
89057..89594
/note="match: GSS: Em:AQ543464"
91625..91990
/note="THE1C repeat: matches 1..371 of consensus"
92079..92237
/note="LIP47 repeat: matches 5949..6125 of consensus"
BASE COUNT 32991 a 18802 c 18896 g 27671 t
ORIGIN

Query Match 1.8%; Score 58; DB 9; Length 98360;
Best Local Similarity 100.0%; Pred. No. 2e-21; 0; Indels 0; Gaps 0;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1187 ctgcccacttgctcccaagtctgggtacagcgtacagcactgcgccagc 1244
|||||
Db 18635 CTGCCACCTTGGCTCCCAAGTCTGGGATACAGCATGACCCACTGCGCCAGC 18578

RESULT 42
AC107939/c
LOCUS AC107939 145178 bp DNA linear HTG 13-FEB-2002
DEFINITION Homo sapiens chromosome 11 clone RP11-4809 map 11, WORKING DRAFT
SEQUENCE, 3 ordered pieces.
ACCESSION AC107939
VERSION AC107939.2 GI:18653699
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 145178)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 11, clone RP11-4809
Unpublished
2 (bases 1 to 145178)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferrelira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Illiev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Larocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C.,
Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., Meldrum,J., Meneus,L., Mihova,T.,
Mlenda,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schnupack,R., Seaman,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,

```

TITLE  
JOURNAL

## COMMENT

zainoun,J., Zembek,L., Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Feb 13, 2002 this sequence version replaced gi:18308666.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information  
Center project name: L24018  
Center clone name: 48\_Q\_9

----- Summary Statistics  
Sequencing vector: Plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 144836 bases at least Q40  
Consensus quality: 144904 bases at least Q30  
Consensus quality: 144952 bases at least Q20  
Insert size: 141000; agarose-fp  
Insert size: 144978; sum-of-contigs  
Quality coverage: 18.7 in Q20 bases; agarose-fp  
Quality coverage: 18.2 in Q20 bases; sum-of-contigs

-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 3 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submittor.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
\* 1 68019: contig of 68019 bp in length  
\* 68020 68119: gap of 100 bp  
\* 68120 140460: contig of 72341 bp in length  
\* 140461 140560: gap of 100 bp  
\* 140561 145178: contig of 4618 bp in length.

## FEATURES

## Source

1..145178  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/map="11"  
/clone="RP11-4809"  
/clone\_lib="RPC1-11 Human Male BAC"

## misc\_feature

1..68019  
/note="assembly\_fragment  
clone\_end:SP6  
vector\_side:left"

## misc\_feature

68120..140460  
/note="assembly\_fragment"

## misc\_feature

140561..145178  
/note="assembly\_fragment  
clone\_end:T7  
vector\_side:right"

BASE COUNT 46626 a 28900 c 28355 g 41097 t 200 others  
ORIGIN

## Query Match

Best Local Similarity 1.8%; Score 58; DB 2; Length 145178;

Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1187 ctgcccacattggctcccaagtctgggtacagcgtacagcactgcgccagc 1244  
|||||

Db 74827 CTGCCACCTTGGCTCCCAAGTCTGGGATACAGCATGACCCACTGCGCCAGC 74770



Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 155982)  
 Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
 Homo sapiens, clone RP11-21E14  
 Unpublished  
 2 (bases 1 to 155982)  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Baldwin, J., Barna, N., Beckerly, R., Boquslavkiy, L., Boukhgalter, B., Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M., Ferrelira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Lehoczy, J., Lieu, C., Locke, K., MacDonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrim, J., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.  
 Direct Submission  
 Submitted (16-NOV-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Sep 10, 2000 this sequence version replaced gi:7249029.  
 All repeats were identified using RepeatMasker:  
 Smit, A. F. A. & Green, P. (1996-1997).  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

TITLE  
 JOURNAL  
 COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: <http://www-seq.wi.mit.edu>  
 Contact: [sequence.submissions@genome.wi.mit.edu](mailto:sequence.submissions@genome.wi.mit.edu)  
 ----- Project Information  
 Center project name: L3991  
 Center clone name: 21\_E14  
 ----- Summary Statistics  
 Sequencing vector: M13; M77815; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.960731  
 Consensus quality: 147683 bases at least Q40  
 Consensus quality: 151981 bases at least Q30  
 Consensus quality: 153418 bases at least Q20  
 Insert size: 156000; agarose-fp  
 Insert size: 154782; sum-of-contigs  
 Quality coverage: 4.4 in Q20 bases; agarose-fp  
 Quality coverage: 4.4 in Q20 bases; sum-of-contigs

\* NOTE: This is a 'working draft' sequence. It currently consists of 13 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 8185: contig of 8185 bp in length  
 \* 186 8285: gap of 100 bp  
 \* 8286 9661: contig of 1376 bp in length  
 \* 9662 9761: gap of 100 bp  
 \* 9762 11131: contig of 1370 bp in length  
 \* 11132 11231: gap of 100 bp  
 \* 11232 57436: contig of 46205 bp in length  
 \* 57437 57536: gap of 100 bp  
 \* 57537 61395: contig of 3859 bp in length  
 \* 61396 61495: gap of 100 bp  
 \* 61496 65614: contig of 4119 bp in length  
 \* 65615 65714: gap of 100 bp  
 \* 65715 71509: contig of 5795 bp in length  
 \* 71510 71609: gap of 100 bp  
 \* 71610 81113: contig of 9504 bp in length  
 \* 81114 81213: gap of 100 bp  
 \* 81214 92173: contig of 10960 bp in length

\* 92174 92273: gap of 100 bp  
 \* 92274 103878: contig of 11605 bp in length  
 \* 103879 103978: gap of 100 bp  
 \* 103979 120091: contig of 16113 bp in length  
 \* 120092 120191: gap of 100 bp  
 \* 120192 154009: contig of 33818 bp in length  
 \* 154010 154109: gap of 100 bp  
 \* 154110 155982: contig of 1873 bp in length.

FEATURES  
 source

1..155982  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="RP11-21E14"  
 /clone\_lib="RPCI-11 Human Male BAC"  
 1..8185  
 /note="assembly\_fragment"  
 clone\_end:SP6  
 vector\_side:left  
 8286..9661  
 /note="assembly\_fragment"  
 9762..11131  
 /note="assembly\_fragment"  
 11232..57436  
 /note="assembly\_fragment"  
 57537..61395  
 /note="assembly\_fragment"  
 61496..65614  
 /note="assembly\_fragment"  
 65715..71509  
 /note="assembly\_fragment"  
 71610..81113  
 /note="assembly\_fragment"  
 81214..92173  
 /note="assembly\_fragment"  
 92274..103878  
 /note="assembly\_fragment"  
 103979..120091  
 /note="assembly\_fragment"  
 120192..154009  
 /note="assembly\_fragment"  
 154110..155982  
 /note="assembly\_fragment"  
 clone\_end:T7  
 vector\_side:right

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

BASE COUNT 49795 a 29060 c 29340 g 46582 t 1205 others  
 ORIGIN

Query Match 1.8%; Score 58; DB 2; Length 155982;  
 Best Local Similarity 100.0%; Pred. No. 2,1e-21;  
 Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1187 ctgccacattggctcccccagtgctgggattacagccatgagccactgcgccagc 1244  
 |||||  
 Db 9997 CTGCCACCTGGCTCCCAAGTGGGATTACAGCATGAGCCACTGCGCCAGC 10054

RESULT 45

AC018988

LOCUS

DEFINITION

IN PROGRESS \*\*\*; 7 unordered pieces.

AC018988

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

Homo sapiens chromosome 15 clone RP11-233C13 map 15, \*\*\* SEQUENCING

157599 bp DNA linear HTG 26-DEC-2001

HTG; HTGS\_PHASE1; HTGS\_FULLTOP; HTGS\_ACTIVEFIN.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 157599)

Birren, B., Linton, L., Nusbaum, C. and Lander, E.

Homo sapiens chromosome 15, clone RP11-233C13

Unpublished

REFERENCE  
AUTHORS

2 (bases 1 to 157599)  
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
 Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Bada,F.,  
 Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,  
 Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,  
 DeRellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,  
 Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,  
 Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,  
 Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,  
 Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K.,  
 Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,  
 Meldrim,J., Meneus,L., Morrow,J., Naylor,J., Norman,C.H.,  
 O'Connor,T., O'Donnell,P., Oliver,T.M., Peterson,K., Pierre,N.,  
 Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D., Roy,A.,  
 Santos,R., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
 Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A.,  
 Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J., Zimmer,A.  
 and Zody,M.

TITLE  
JOURNAL  
COMMENT

Direct Submission  
 Submitted (25-DEC-1999) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Dec 26, 2001 this sequence version replaced gi:14010824.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 ----- Project Information  
 Center project name: L5451  
 Center clone name: 233\_C\_13  
 -----

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 7 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

\* 1 30289: contig of 30289 bp in length  
 \* 30290 30389: gap of 100 bp  
 \* 30390 41503: contig of 11114 bp in length  
 \* 41504 41603: gap of 100 bp  
 \* 41604 107568: contig of 65965 bp in length  
 \* 107569 107668: gap of 100 bp  
 \* 107669 119412: contig of 11744 bp in length  
 \* 119413 119512: gap of 100 bp  
 \* 119513 139173: contig of 19661 bp in length  
 \* 139174 139273: gap of 100 bp  
 \* 139274 149978: contig of 10705 bp in length  
 \* 149979 150078: gap of 100 bp  
 \* 150079 157599: contig of 7521 bp in length.

FEATURES  
source

1..157599  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /map="15"  
 /clone="RP11-233C13"  
 /clone\_lib="RPC1-11 Human Male BAC"  
 BASE COUNT 41693 a 38435 c 38160 g 38669 t 642 others  
 ORIGIN

Query Match 1.8%; Score 58; DB 2; Length 157599;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-21;  
 Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1187 ctgccacattggcctcccaagtctgggattacaggcatgagccactgcccagc 1244  
 |||

Db 145556 CTGCCACCTTTGGCCTCCCAAAGTGTGGGATTACAGGCATGAGCCACTGGGCCAGC 145613

Search completed: September 20, 2002, 06:25:43  
 Job time: 18137 sec







GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.  
OM nucleic - nucleic search, using sw model  
Run on: September 20, 2002, 03:15:16 ; Search time 521.76 Seconds  
(without alignments)  
10631.997 Million cell updates/sec

Title: US-09-846-456-1  
Perfect score: 3231  
Sequence: 1 acagggcatgtgtgcagggtg.....gccccacatccccaccattt 3231  
Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0  
Searched: 1736436 seqs, 858457221 residues  
Word size : 0  
Total number of hits satisfying chosen parameters: 3472872  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Listing first 45 summaries

Database :			
N_Geneseq_032802:*			
1:	/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1980.DAT:*		
2:	/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1981.DAT:*		
3:	/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1982.DAT:*		
4:	/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1983.DAT:*		
5:	/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1984.DAT:*		
6:	/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1985.DAT:*		
7:	/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1986.DAT:*		
8:	/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1987.DAT:*		
9:	/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1988.DAT:*		
10:	/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1989.DAT:*		
11:	/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1990.DAT:*		
12:	/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1991.DAT:*		
13:	/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1992.DAT:*		
14:	/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1993.DAT:*		
15:	/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1994.DAT:*		
16:	/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT:*		
17:	/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1996.DAT:*		
18:	/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1997.DAT:*		
19:	/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1998.DAT:*		
20:	/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1999.DAT:*		
21:	/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:*		
22:	/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:*		
23:	/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT:*		
24:	/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT:*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length DB ID	Description
1	814	25.2 183999 22	AAF92831 Human ABC1 genomic
2	541	16.7 1643 22	AAF24681 Nucleotide sequenc
3	541	16.7 1643 22	AAF24703 Nucleotide sequenc
4	336	10.4 763 22	AAH04729 Human cDNA clone (
5	336	10.4 1750 22	AAH17451 Human cDNA sequenc
6	228	7.1 7260 22	AAH21326 Human ATP binding
7	228	7.1 7260 22	AAH170315 Human ATP binding
8	217	6.7 736 22	AAH07432 Human cDNA clone (
9	217	6.7 1556 22	AAH18606 Human cDNA sequenc

10	210	6.5	227	21	AAC09615	Human secreted pro
11	205	6.3	7086	22	ABA09200	Human ABCA1 homolo
12	205	6.3	7086	22	AAK52667	Human polynucleoti
13	205	6.3	7281	22	AAK51683	Human polynucleoti
14	205	6.3	9854	22	AAS06121	Human ABC1 DNA seq
15	197	6.1	10442	22	AAF24680	Nucleotide sequenc
16	197	6.1	10442	22	AAF24702	Nucleotide sequenc
17	188	5.8	10474	22	AAF24685	Nucleotide sequenc
18	188	5.8	10474	22	AAF24686	Nucleotide sequenc
19	188	5.8	10474	22	AAF24707	Nucleotide sequenc
20	188	5.8	10474	22	AAF24708	Nucleotide sequenc
21	92	2.8	446	22	AAS04035	Partial human ABC1
22	92	2.8	9741	22	AAS06120	Human ABC1 DNA seq
23	55	1.7	235033	19	AAV57926	Hereditary haemoch
24	55	1.7	237326	19	AAV57903	Hereditary haemoch
25	52	1.6	1316	22	AAI64613	Human ribosomal S1
26	52	1.6	5351	22	AAS31466	Human DNA for a no
27	52	1.6	6461	22	AAS31467	Human immune/haema
28	52	1.6	8319	22	AAK65197	Human immunorep
29	52	1.6	10901	22	AAI03236	Human nervous syst
30	52	1.6	29329	22	ABA18026	Human immune/haema
31	52	1.6	29329	22	ABA20511	Human immune/haema
32	52	1.6	29329	22	AAK70791	Human secreted pro
33	52	1.6	29329	22	AAK78512	Human polynucleoti
34	51	1.6	149	21	AAK25949	Human polynucleoti
35	51	1.6	396	22	AAI81653	Human immune/haema
36	51	1.6	425	22	AAI87958	Human immune/haema
37	51	1.6	1856	22	AAK77921	Human polypeptide-
38	51	1.6	2046	22	AAI64545	Human immune/haema
39	51	1.6	7759	22	AAK77916	Human ovarian and
40	51	1.6	17904	22	ABA07913	Human reproductive
41	51	1.6	17904	22	AAI03730	Human musculoskele
42	51	1.6	26591	22	AAI36313	Human DNA for a no
43	51	1.6	32186	22	AAS34422	Human thioredoxin
44	51	1.6	66566	21	AAAS3450	Human histone deac
45	51	1.6	122186	22	AAC89560	

ALIGNMENTS

RESULT	1
AAF92831	
ID	AAF92831 standard; DNA; 183999 BP.
XX	
AC	AAF92831;
XX	
DT	17-MAY-2001 (first entry)
XX	
DE	Human ABC1 genomic DNA.
XX	
KW	High density lipoprotein-cholesterol; HDL-C; cardiovascular; ABC1; ds.
XX	
OS	Homo sapiens.
XX	
PN	WO200115676-A2.
XX	
PD	08-MAR-2001.
XX	
PF	01-SEP-2000; 2000WO-IB01492.
XX	
PR	01-SEP-1999; 99US-0151977.
PR	15-MAR-2000; 2000US-0526193.
PR	23-JUN-2000; 2000US-0213958.
XX	
PA	(UYBR-) UNIV BRITISH COLUMBIA.
PA	(XENO-) XENON GENETICS INC.
XX	
PI	Hayden MR, Brooks-Wilson AR, Pimstone SN, Clee SM;
XX	
DR	WPI; 2001-244356/25.
XX	
PT	Treating a lower than normal high density lipoprotein-cholesterol

PT (HDL-C) level, a higher than normal triglyceride level, or a  
 PT cardiovascular disease, by administering a compound that modulates LXR-  
 PT or RXR-mediated transcriptional activity -  
 PS Claim 8; Fig 1; 317pp; English.

XX The present invention relates to a method for treating a patient  
 CC diagnosed as having a lower than normal high density  
 CC lipoprotein-cholesterol (HDL-C) level, a higher than normal  
 CC triglyceride level, or a cardiovascular disease, involving  
 CC administering a compound that modulates LXR- or RXR-mediated  
 CC transcriptional activity or ABC1 expression or identity.  
 CC The LXR gene product may be used in an assay to identify  
 CC compounds useful for the treatment of a disease or condition selected a  
 CC lower than normal HDL cholesterol level, a higher than normal  
 CC triglyceride level, and a cardiovascular disease.

XX Sequence 183999 BP; 49549 A; 37944 C; 41170 G; 54950 T; 386 other;

Query Match 25.2%; Score 814; DB 22; Length 183999;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 964; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 782 atgaacagagggcagaaagactttacgttaaaattgctcattggtgttcaagtttgac 841  
 DB 26671 atgaacagagggcagaaagactttacgttaaaattgctcattggtgttcaagtttgac 26730  
 QY 842 ccacaaacccaattttatgacaaaggtttattcttgaactgagcaagggtccgctctc 901  
 DB 26731 ccacaaacccaattttatgacaaaggtttattcttgaactgagcaagggtccgctctc 26790  
 QY 902 ctgggcttgggtttagaagctcatctctggcctttctgagatccatcccttttttt 961  
 DB 26791 ctgggcttgggtttagaagctcatctctggcctttctgagatccatcccttttttt 26850  
 QY 962 tattttctgacacgaggttctctctcactcaagctgagtgagtgagtgagtgagtgag 1021  
 DB 26851 tattttctgacacgaggttctctctcactcaagctgagtgagtgagtgagtgagtgag 26910  
 QY 1022 cgactcaactgaactctgctcccggttcaagcgattctcctcagcctcctgag 1081  
 DB 26911 cgactcaactgaactctgctcccggttcaagcgattctcctcagcctcctgag 26970  
 QY 1082 ataacagggccgcgcacacacatctgctcaattttgtttatgttttaaaagactgggtt 1141  
 DB 26971 ataacagggccgcgcacacacatctgctcaattttgtttatgttttaaaagactgggtt 27030  
 QY 1142 catcatgttggccaggttgggtttcgaactcctgacctgagtgagtgagtgagtgagtgag 1201  
 DB 27031 catcatgttggccaggttgggtttcgaactcctgacctgagtgagtgagtgagtgagtgag 27090  
 QY 1202 tcccaaatgtcgggattacagagctgagcactgcgccagctcagatccatcccttc 1261  
 DB 27091 tcccaaatgtcgggattacagagctgagcactgcgccagctcagatccatcccttc 27150  
 QY 1262 taaggcacaacagtcacatgttgtaaaagggggccatgccaccagagttatgagtaactggg 1321  
 DB 27151 taaggcacaacagtcacatgttgtaaaagggggccatgccaccagagttatgagtaactggg 27210  
 QY 1322 actcagaattccttgcttgggttgcacatgcacttccaggcctgcttgggctctc 1381  
 DB 27211 actcagaattccttgcttgggttgcacatgcacttccaggcctgcttgggctctc 27270  
 QY 1382 tctatgctctgctgagtgatgtgtatagaaccactgatgtgagtaactgggcttgagcgg 1441  
 DB 27271 tctatgctctgctgagtgatgtgtatagaaccactgatgtgagtaactgggcttgagcgg 27330  
 QY 1442 tggcctgagatccctgttgaactgagcatgagggggccttgcagactgaatgtctgcat 1501  
 DB 27331 tggcctgagatccctgttgaactgagcatgagggggccttgcagactgaatgtctgcat 27390  
 QY 1502 gcaggttgggagttcttggaaatgatcgtgagctggaggtgggaagagaagtaggcttgg 1561

DB 27391 gcaggttgggagttcttggaaatgatcgtgagctggaggtgggaagagaagtaggcttgg 27450  
 QY 1562 ggcagctctctcatgccacctcattctggccaaactcaggtcacaactgtgaagagtcta 1621  
 DB 27451 ggcagctctctcatgccacctcattctggccaaactcaggtcacaactgtgaagagtcta 27510  
 QY 1622 aatgtgaatctccctcaagggtggtcactcaaaagggtatcttctgcaaggtaggagaccttg 1681  
 DB 27511 aatgtgaatctccctcaagggtggtcactcaaaagggtatcttctgcaaggtaggagaccttg 27570  
 QY 1682 tggcctccacgtgcacttcccaagggtcctgttggcctcttctacgggtctgctccctgagtc 1741  
 DB 27571 tggcctccacgtgcacttcccaagggtcctgttggcctcttctacgggtctgctccctgagtc 27630  
 QY 1742 ttctatg 1748  
 DB 27631 ttctatg 27637

RESULT 2  
 AAF24681  
 ID AAF24681 standard; DNA; 1643 BP.

XX AAF24681;

DT 20-APR-2001 (first entry)

XX Nucleotide sequence of the 5' flanking region of the human ABC1 gene.

DE Human; adenosine triphosphate binding cassette protein 1; ABC1;  
 KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;  
 KW chromosome 9q22-q31; heart disease; hypercholesterolemia;  
 KW atherosclerosis; cholesterol transport; ss.

OS Homo sapiens.

XX WO200078972-A2.

PN 28-DEC-2000.

PD 16-JUN-2000; 2000WO-US16765.

XX 18-JUN-1999; 99US-0140264.

PR 14-SEP-1999; 99US-0153872.

PR 19-NOV-1999; 99US-0166573.

XX (CVTH-) CV THERAPEUTICS INC.

PA Lawn RM, Wade D, Garvin M;

XX WPI; 2001-137812/14.

XX Adenosine triphosphate (ATP) binding cassette (ABC) polynucleotide,  
 useful for the development of agents for the treatment of heart disease  
 and other disorders associated with hypercholesterolemia and  
 atherosclerosis -

Claim 1; Page 143-144; 215pp; English.

XX The present sequence represents the 5' flanking region of the human  
 CC adenosine triphosphate (ATP) binding cassette protein (ABC) 1 gene. ABC1  
 CC resides in cell membranes and utilizes ATP hydrolysis to transport a wide  
 CC variety of substrates across the plasma membrane. ABC1 is a pivotal  
 CC protein in the apolipoprotein-mediated mobilisation of intracellular  
 CC cholesterol stores. ABC1 is defective in Tangier disease, a genetic  
 CC disorder characterised by abnormal HDL-cholesterol metabolism. The ABC1  
 CC gene is localised to chromosome 9q22-q31. The ABC1 genes and proteins  
 CC are useful for developing pharmaceutical agents for the treatment of  
 CC heart disease and other disorders associated with hypercholesterolemia  
 CC and atherosclerosis. The genes are useful for developing screening assays  
 CC to screen for compounds that regulate the expression of genes associated  
 CC with cholesterol transport. The genes and proteins are also useful for



Qy 2610 gaagggagcagaccgagaccctaagacacactgtgtaccctccaccaccaccacc 2669  
 Db 1274 gaagggagcagaccgagaccctaagacacactgtgtaccctccaccaccaccacc 1333  
 Qy 2670 c 2670  
 Db 1334 c 1334

RESULT 4  
 AAH04729  
 ID AAH04729 standard; cDNA; 763 BP.  
 AC AAH04729;  
 XX  
 XX 26-JUN-2001 (first entry)  
 DT  
 XX  
 DE Human cDNA clone (5'-primer) SEQ ID NO:1564.  
 XX  
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP1074617-A2.  
 XX  
 PD 07-FEB-2001.  
 XX  
 XX 28-JUL-2000; 2000EP-0116126.  
 XX  
 PR 29-JUL-1999; 99JP-0248036.  
 PR 27-AUG-1999; 99JP-0300253.  
 PR 11-JAN-2000; 2000JP-0118776.  
 PR 02-MAY-2000; 2000JP-0183767.  
 PR 09-JUN-2000; 2000JP-0241899.  
 XX  
 PA (HELI-) HELIX RES INST.  
 XX  
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX  
 DR WPI; 2001-318749/34.  
 XX  
 XX Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs -  
 XX  
 PS Claim 1; SEQ ID 1564; 2537pp + CD ROM; English.  
 XX  
 CC The present invention describes primer sets for synthesizing 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to  
 CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.

SQ Sequence 763 BP; 137 A; 205 C; 260 G; 158 T; 3 other;

Query Match 10.4%; Score 336; DB 22; Length 763;  
 Best Local Similarity 100.0%; Pred. No. 4.1e-143;  
 Matches 336; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2896 aattgcgagcagagtgagtgaggcgccgagcccgagagccgagcccttctctcc 2955  
 Db 1 aattgcgagcagagtgagtgaggcgccgagcccgagagccgagcccttctctcc 60  
 QY 2956 gggctgcggcagggcagggcggggagctccgcgaccacagagccggttctcaggcgcc 3015  
 Db 61 gggctgcggcagggcagggcggggagctccgcgaccacagagccggttctcaggcgcc 120  
 QY 3016 ttgtctctgttttttcccggttctgttttctcccttctccggaaggttctcaagg 3075  
 Db 121 ttgtctctgttttttcccggttctgttttctcccttctccggaaggttctcaagg 180  
 QY 3076 ggtagagaaaagagcgcgcaaacacacaaagtggaaacaggttaagaggtctctcagtgact 3135  
 Db 181 ggtagagaaaagagcgcgcaaacacacaaagtggaaacaggttaagaggtctctcagtgact 240  
 QY 3136 tacttggcggttattgtttgttgcgagccaaagagcgttcggaaggtctcggtttcg 3195  
 Db 241 tacttggcggttattgtttgttgcgagccaaagagcgttcggaaggtctcggtttcg 300  
 QY 3196 gggactttgatccgagccacacatccccaccatt 3231  
 Db 301 gggactttgatccgagccacacatccccaccatt 336

## RESULT 5

AAH17451  
 ID AAH17451 standard; cDNA; 1750 BP.

XX AC AAH17451;  
 XX

XX 26-JUN-2001 (first entry)  
 DT

XX Human cDNA sequence SEQ ID NO:16905.  
 DE

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
 KW

XX Homo sapiens.  
 OS

XX EP1074617-A2.  
 PN

XX 07-FEB-2001.  
 PD

XX 28-JUL-2000; 2000EP-0116126.  
 XX

XX 29-JUL-1999; 99JP-0248036.  
 PR

XX 27-AUG-1999; 99JP-0300253.  
 PR

XX 11-JAN-2000; 2000JP-0118776.  
 PR

XX 02-MAY-2000; 2000JP-0183767.  
 PR

XX 09-JUN-2000; 2000JP-0241899.  
 XX

XX (HELI-) HELIX RES INST.  
 PA

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 PI

XX WPI; 2001-318749/34.  
 DR

XX Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs -

XX Claim 8; SEQ ID 16905; 2537pp + CD ROM; English.  
 PS

XX The present invention describes primer sets for synthesizing 5602  
 CC







CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.  
 XX  
 SQ Sequence 736 BP; 163 A; 199 C; 199 G; 170 T; 5 other;

Query Match 6.7%; Score 217; DB 22; Length 736;  
 Best Local Similarity 100.0%; Pred. No. 8.1e-89;  
 Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2898 ttgcgagcagagtgagtgaggccggaccgcagagccgagccgaccttctctccgg 2957  
 Db 5 ttgcgagcagagtgagtgaggccggaccgcagagccgagccgaccttctctccgg 64  
 QY 2958 gctgcggcaggcaggcgggagctccgcgaccacagagccggttctcaggcgctt 3017  
 Db 65 gctgcggcaggcaggcgggagctccgcgaccacagagccggttctcaggcgctt 124  
 QY 3018 tgctcctgtttttcccggttctgtttctccttctccgaaaggcttcaaggcg 3077  
 Db 125 tgctcctgtttttcccggttctgtttctccttctccgaaaggcttcaaggcg 184  
 QY 3078 taggagaaagagcagcaacacaaagtggaaacag 3114  
 Db 185 taggagaaagagcagcaacacaaagtggaaacag 221

RESULT 9  
 AAH18606  
 ID AAH18606 standard; cDNA; 1556 BP.  
 XX  
 AC AAH18606;  
 XX  
 XX  
 XX 26-JUN-2001 (first entry)  
 XX  
 DE Human cDNA sequence SEQ ID NO:18808.  
 XX  
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
 KW Homo sapiens.  
 XX  
 XX EPI074617-A2.  
 XX  
 XX 07-FEB-2001.  
 XX  
 XX 28-JUL-2000; 2000EP-0116126.  
 XX  
 XX 29-JUL-1999; 99JP-0248036.  
 XX 27-AUG-1999; 99JP-0300253.  
 XX 11-JAN-2000; 2000JP-0118776.  
 XX 02-MAY-2000; 2000JP-0183767.  
 XX 09-JUN-2000; 2000JP-0241899.  
 XX  
 PA (HELI-) HELIX RES INST.  
 XX  
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX  
 XX WPI; 2001-318749/34.  
 XX  
 XX Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs -  
 XX  
 PS Claim 8; SEQ ID 18808; 2537pp + CD ROM; English.  
 XX

CC The present invention describes primer sets for synthesising 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC in gene therapy. The primers are useful for synthesising polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.  
 XX  
 SQ Sequence 1556 BP; 380 A; 363 C; 399 G; 414 T; 0 other;

Query Match 6.7%; Score 217; DB 22; Length 1556;  
 Best Local Similarity 100.0%; Pred. No. 8.1e-89;  
 Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2898 ttgcgagcagagtgagtgaggccggaccgcagagccgagccgaccttctctccgg 2957  
 Db 5 ttgcgagcagagtgagtgaggccggaccgcagagccgagccgaccttctctccgg 64  
 QY 2958 gctgcggcaggcaggcgggagctccgcgaccacagagccggttctcaggcgctt 3017  
 Db 65 gctgcggcaggcaggcgggagctccgcgaccacagagccggttctcaggcgctt 124  
 QY 3018 tgctcctgtttttcccggttctgtttctccttctccgaaaggcttcaaggcg 3077  
 Db 125 tgctcctgtttttcccggttctgtttctccttctccgaaaggcttcaaggcg 184  
 QY 3078 taggagaaagagcagcaacacaaagtggaaacag 3114  
 Db 185 taggagaaagagcagcaacacaaagtggaaacag 221

RESULT 10  
 AAC09615  
 ID AAC09615 standard; cDNA; 227 BP.  
 XX  
 AC AAC09615;  
 XX  
 XX 06-OCT-2000 (first entry)  
 DT  
 XX Human secreted protein 5' EST, SEQ ID NO: 13690.  
 DE  
 XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
 KW gene therapy; chromosome mapping; ss.  
 KW Homo sapiens.  
 XX  
 XX EP1033401-A2.  
 PN  
 XX 06-SEP-2000.  
 PD  
 XX 21-FEB-2000; 2000EP-0200610.  
 PF  
 XX 26-FEB-1999; 99US-0122487.  
 PR  
 XX (GEST ) GENSET.  
 PA  
 XX Dumas Milne Edwards J, Duclert A, Giordano J;  
 PI



|||||  
Db 67 caggcgccgagctccgcgacacagagccggtcttcaggcgcttctgctctctttt 126  
QY 3030 ttcccccgttctgttttctcccttctccggaaggcttgcgaagggttaggagaaagag 3089  
Db 127 ttcccccgttctgttttctcccttctccggaaggcttgcgaagggttaggagaaagag 186  
QY 3090 acgcaaacacaaaagtgaataacag 3114  
Db 187 acgcaaacacaaaagtgaataacag 211  
RESULT 12  
AAK52667  
ID AAK52667 standard; cDNA; 7086 BP.  
XX  
AC AAK52667;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human polynucleotide SEQ ID NO 2196.  
XX  
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorder; arthritis; inflammation; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200157190-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 05-FEB-2001; 2001WO-US04098.  
XX  
PR 03-FEB-2000; 2000US-0496914.  
PR 27-APR-2000; 2000US-0560875.  
PR 20-JUN-2000; 2000US-0598075.  
PR 19-JUL-2000; 2000US-0620325.  
PR 01-SEP-2000; 2000US-0654936.  
PR 15-SEP-2000; 2000US-0663561.  
PR 20-OCT-2000; 2000US-0693325.  
PR 30-NOV-2000; 2000US-0728422.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;  
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;  
XX  
DR WPI: 2001-476283/51.  
DR P-PSDB; AAM79534.  
XX  
PT Nucleic acids encoding polypeptides with cytokine-like activities,  
PT useful in diagnosis and gene therapy -  
XX  
PS Claim 1; Page 4558-4560; 6221pp; English.  
XX  
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the  
CC encoded polypeptides (AAM78323-AAAM80302) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation.  
CC  
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666  
CC (AAM80020) are omitted as the relevant pages from the sequence listing  
CC were missing at the time of publication.  
XX

SQ Sequence 7086 BP; 1773 A; 1739 C; 1859 G; 1715 T; 0 other;  
Query Match 6.3%; Score 205; DB 22; Length 7086;  
Best Local Similarity 100.0%; Pred. NO. 2.4e-83;  
Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2910 gtgagtgccggccgacccgcagagccgaccccttctcccggtcgcgcagag 2969  
Db 7 gtgagtgccggccgacccgcagagccgaccccttctcccggtcgcgcagag 66  
QY 2970 caggcgccggagctccgcgcacacacagagccggttctcaggcgcttctgctctgttt 3029  
Db 67 caggcgccggagctccgcgcacacacagagccggttctcaggcgcttctgctctgttt 126  
QY 3030 ttcccccgttctgttttctcccttctccggaaggcttgcgaagggttaggagaaagag 3089  
Db 127 ttcccccgttctgttttctcccttctccggaaggcttgcgaagggttaggagaaagag 186  
QY 3090 acgcaaacacaaaagtgaataacag 3114  
Db 187 acgcaaacacaaaagtgaataacag 211  
RESULT 13  
AAK51683  
ID AAK51683 standard; cDNA; 7281 BP.  
XX  
AC AAK51683;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human polynucleotide SEQ ID NO 228.  
XX  
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorder; arthritis; inflammation; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200157190-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 05-FEB-2001; 2001WO-US04098.  
XX  
PR 03-FEB-2000; 2000US-0496914.  
PR 27-APR-2000; 2000US-0560875.  
PR 20-JUN-2000; 2000US-0598075.  
PR 19-JUL-2000; 2000US-0620325.  
PR 01-SEP-2000; 2000US-0654936.  
PR 15-SEP-2000; 2000US-0663561.  
PR 20-OCT-2000; 2000US-0693325.  
PR 30-NOV-2000; 2000US-0728422.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;  
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;  
XX  
DR WPI: 2001-476283/51.  
DR P-PSDB; AAM78550.  
XX  
PT Nucleic acids encoding polypeptides with cytokine-like activities,  
PT useful in diagnosis and gene therapy -  
XX  
PS Claim 1; Page 1086-1096; 6221pp; English.  
XX  
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the  
CC encoded polypeptides (AAM78323-AAAM80302) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce

CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activin/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation.  
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666  
 CC (AAM80020) are omitted as the relevant pages from the sequence listing  
 CC were missing at the time of publication.  
 XX  
 SQ Sequence 7281 BP; 1831 A; 1773 C; 1915 G; 1762 T; 0 other;

Query Match 6.3%; Score 205; DB 22; Length 7281;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-83;  
 Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2910 gtgagtggcgccggaccgcagagccgagccgagccacccttctcccggtcgccgaggg 2969  
 |||||||  
 Db 45 gtgagtggcgccggaccgcagagccgagccgagccacccttctcccggtcgccgaggg 104  
 QY 2970 caggcgccggagctccgcgaccacacagaccggttctcaggcgcttgcctctgttt 3029  
 |||||||  
 Db 105 caggcgccggagctccgcgaccacacagaccggttctcaggcgcttgcctctgttt 164  
 QY 3030 ttcccggttctgtttctcccttctccggaaggctgtcgaaggtaggagaaagag 3089  
 |||||||  
 Db 165 ttcccggttctgtttctcccttctccggaaggctgtcgaaggtaggagaaagag 224  
 QY 3090 acgcaaacacaaaagtggaaaacag 3114  
 |||||||  
 Db 225 acgcaaacacaaaagtggaaaacag 249

## RESULT 14

AA06121  
 ID AAS06121 standard; cDNA; 9854 BP.

XX AC AAS06121;

XX 12-SEP-2001 (first entry)

XX Human ABC1 DNA sequence #2.

XX Human; ABC1 gene; atherosclerosis; reverse transport; cholesterol;  
 KW cardiovascular; neurological; Tangier disease; LCAT deficiency;  
 KW lecithin-cholesterol acetyltransferase; malaria; diabetes; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 298..7078  
 FT /\*tag= a  
 FT /product= "Human ABC1 protein"

PN WO200130848-A2.

XX 03-MAY-2001.

XX 26-OCT-2000; 2000WO-EF10886.

XX 26-OCT-1999; 99EP-0402668.

PR 01-MAR-2000; 2000US-0186260.

XX (AVET ) AVENTIS PHARMA SA.

XX Denefle P, Rosier-Montus M, Arnould-Requigne I, Prades C, Naudin L;  
 PI Lemoine C, Duverger N, Jaye M, Searfoss GH, Remaley A, Brewer HB;  
 PI Dean M;

XX WPI; 2001-316327/33.

DR

DR P-PSDB; AAU02176.

XX New human ABC1 nucleic acids and polypeptides for treating  
 PT atherosclerosis, malaria and diabetes -

XX Claim 1; Page 209-213; 368pp; English.

CC The sequence represents the coding sequence #2 of human ABC1. The  
 CC nucleic acid sequence, primers and probes derived from the ABC1 sequence,  
 CC and polypeptides and vectors are useful for the prevention of  
 CC atherosclerosis, in a subject affected by a dysfunction in the reverse  
 CC transport of cholesterol. The polypeptide encoded by the ABC1 gene is  
 CC useful for screening for an active ingredient for the prevention or  
 CC treatment of a disease resulting from dysfunction in the reverse  
 CC transport of cholesterol. The nucleic acids and polypeptides are also  
 CC useful for treating and preventing cardiovascular and neurological  
 CC pathologies, and other diseases e.g. Tangier disease, lecithin-  
 CC cholesterol (LCAT) deficiency, malaria and diabetes.

XX Sequence 9854 BP; 2665 A; 2219 C; 2334 G; 2635 T; 1 other;

Query Match 6.3%; Score 205; DB 22; Length 9854;

Best Local Similarity 100.0%; Pred. No. 2.4e-83;

Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2910 gtgagtggcgccggaccgcagagccgagccgagccacccttctcccggtcgccgaggg 2969  
 |||||||  
 Db 1 gtgagtggcgccggaccgcagagccgagccgagccacccttctcccggtcgccgaggg 60

QY 2970 caggcgccggagctccgcgaccacacagaccggttctcaggcgcttgcctctgttt 3029  
 |||||||  
 Db 61 caggcgccggagctccgcgaccacacagaccggttctcaggcgcttgcctctgttt 120

QY 3030 ttcccggttctgtttctcccttctccggaaggctgtcgaaggtaggagaaagag 3089  
 |||||||  
 Db 121 ttcccggttctgtttctcccttctccggaaggctgtcgaaggtaggagaaagag 180

QY 3090 acgcaaacacaaaagtggaaaacag 3114  
 |||||||

Db 181 acgcaaacacaaaagtggaaaacag 205

## RESULT 15

AAF24680  
 ID AAF24680 standard; DNA; 10442 BP.

XX AC AAF24680;

XX 20-APR-2001 (first entry)

XX Nucleotide sequence of a human ABC1 polypeptide.

XX Human; adenosine triphosphate binding cassette protein 1; ABC1;  
 KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;  
 KW chromosome 9q22-9q31; heart disease; hypercholesterolemia;  
 KW atherosclerosis; cholesterol transport; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 291..7076  
 FT /\*tag= a  
 FT /product= "ABC1 polypeptide"

PN WO200078972-A2.

XX 28-DEC-2000.

XX 16-JUN-2000; 2000WO-US16765.

XX 18-JUN-1999; 99US-0140264.

PR 14-SEP-1999; 99US-0153872.

```

PR 19-NOV-1999; 99US-0166573.
XX (CVTH-) CV THERAPEUTICS INC.
XX PA
XX PI Lawn RM, Wade D, Garvin M;
XX DR WPI; 2001-137812/14.
XX XX
XX Adenosine triphosphate (ATP) binding cassette (ABC) polynucleotide,
XX useful for the development of agents for the treatment of heart disease
XX and other disorders associated with hypercholesterolemia and
XX atherosclerosis -
XX PS Disclosure; Page 122-128; 215pp; English.
XX CC
XX The present sequence encodes a human adenosine triphosphate (ATP)
XX binding cassette protein (ABC) 1 polypeptide. ABC1 resides in cell
XX membranes and utilises ATP hydrolysis to transport a wide variety of
XX substrates across the plasma membrane. ABC1 is a pivotal protein in
XX the apolipoprotein-mediated mobilisation of intracellular cholesterol
XX stores. ABC1 is defective in Tangier disease, a genetic disorder
XX characterised by abnormal HDL-cholesterol metabolism. The ABC1 gene is
XX localised to chromosome 9q22-9q31. The ABC1 genes and proteins are
XX useful for developing pharmaceutical agents for the treatment of heart
XX disease and other disorders associated with hypercholesterolemia and
XX atherosclerosis. The genes are useful for developing screening assays to
XX screen for compounds that regulate the expression of genes associated
XX with cholesterol transport. The genes and proteins are also useful for
XX CC other disorders associated with hypercholesterolemia.
XX SQ Sequence 10442 BP; 2898 A; 2297 C; 2408 G; 2835 T; 4 other;

Query Match 6.1%; Score 197; DB 22; Length 10442;
Best Local Similarity 100.0%; Pred. No. 1.1e-79;
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2918 ggccgggaccgcagagccgagccgacccctctctccgggctgcggcaggcaggcg 2977
Db 1 ggccgggaccgcagagccgagccgacccctctctccgggctgcggcaggcaggcg 60
QY 2978 ggaagtcgcgcacacacagcggcttcacaggcgcttgcctgtttttcccg 3037
Db 61 ggaagtcgcgcacacacagcggcttcacaggcgcttgcctgtttttcccg 120
QY 3038 gttctgttttcccttccggaaggctgtcagggttaggagaagacgcaaac 3097
Db 121 gttctgttttcccttccggaaggctgtcagggttaggagaagacgcaaac 180
QY 3098 acaaaagtggaaaacag 3114
Db 181 acaaaagtggaaaacag 197

RESULT 16
AAF24702
ID AAF24702 standard; DNA; 10442 BP.
XX AC
XX AAF24702;
XX XX
XX 20-APR-2001 (first entry)
XX DE
XX Nucleotide sequence of a human ABC1 polypeptide.
XX KW Human; adenosine triphosphate binding cassette protein 1; ABC1;
XX apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;
XX KW chromosome 9q22-9q31; heart disease; hypercholesterolemia;
XX KW atherosclerosis; cholesterol transport; ss.
XX XX
XX Homo sapiens.
XX OS
XX XX
XX Key Location/Qualifiers

```

---

```

FT CDS 291..7076
FT /*tag= a
XX /product= "ABC1 polypeptide"
XX PN WO200078971-A2.
XX PD 28-DEC-2000.
XX XX
XX 16-JUN-2000; 2000WO-US16591.
XX PR 18-JUN-1999; 99US-0140264.
XX PR 14-SEP-1999; 99US-0153872.
XX PR 19-NOV-1999; 99US-0166573.
XX XX
XX (CVTH-) CV THERAPEUTICS INC.
XX PA (UNIW ) UNIV WASHINGTON.
XX XX
XX PI Lawn RM, Wade D, Oram JF, Garvin M;
XX XX
XX WPI; 2001-137811/14.
XX DR P-PSDB; AAB31365.
XX XX
XX Adenosine triphosphate (ATP) binding cassette protein (ABC) 1
XX polynucleotides and polypeptides, useful for treatment of heart disease
XX PT and other disorders associated with hypercholesterolemia and
XX PT atherosclerosis -
XX PT
XX PT
XX Claim 3; Page 117-123; 211pp; English.
XX CC
XX The present sequence encodes a human adenosine triphosphate (ATP)
XX binding cassette protein (ABC) 1 polypeptide. ABC1 resides in cell
XX membranes and utilises ATP hydrolysis to transport a wide variety of
XX substrates across the plasma membrane. ABC1 is a pivotal protein in
XX the apolipoprotein-mediated mobilisation of intracellular cholesterol
XX stores. ABC1 is defective in Tangier disease, a genetic disorder
XX characterised by abnormal HDL-cholesterol metabolism. The ABC1 gene is
XX localised to chromosome 9q22-9q31. The ABC1 genes and proteins are
XX useful for developing pharmaceutical agents for the treatment of heart
XX disease and other disorders associated with hypercholesterolemia and
XX atherosclerosis. The genes are useful for developing screening assays to
XX screen for compounds that regulate the expression of genes associated
XX with cholesterol transport. The genes and proteins are also useful for
XX CC other disorders associated with hypercholesterolemia.
XX XX
XX Sequence 10442 BP; 2898 A; 2297 C; 2408 G; 2835 T; 4 other;

Query Match 6.1%; Score 197; DB 22; Length 10442;
Best Local Similarity 100.0%; Pred. No. 1.1e-79;
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2918 ggccgggaccgcagagccgagccgacccctctctccgggctgcggcaggcaggcg 2977
Db 1 ggccgggaccgcagagccgagccgacccctctctccgggctgcggcaggcaggcg 60
QY 2978 ggaagtcgcgcacacacagcggcttcacaggcgcttgcctgtttttcccg 3037
Db 61 ggaagtcgcgcacacacagcggcttcacaggcgcttgcctgtttttcccg 120
QY 3038 gttctgttttcccttccggaaggctgtcagggttaggagaagacgcaaac 3097
Db 121 gttctgttttcccttccggaaggctgtcagggttaggagaagacgcaaac 180
QY 3098 acaaaagtggaaaacag 3114
Db 181 acaaaagtggaaaacag 197

RESULT 17
AAF24685
ID AAF24685 standard; DNA; 10474 BP.
XX XX

```

AC AAF24685;  
 XX 20-APR-2001 (first entry)  
 XX Nucleotide sequence of ABC1 polypeptide from Tangier disease patient.  
 DE Human; adenosine triphosphate binding cassette protein 1; ABC1;  
 XX apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;  
 KW chromosome 9q22-q31; heart disease; hypercholesterolemia;  
 KW atherosclerosis; cholesterol transport; ss.  
 XX Homo sapiens.  
 OS  
 XX  
 XX Key Location/Qualifiers  
 FH 323..7108  
 FT /\*tag= a  
 FT /product= "defective ABC1 polypeptide"  
 XX  
 XX WO200078972-A2.  
 XX  
 XX 28-DEC-2000.  
 XX  
 XX 16-JUN-2000; 2000WO-US16765.  
 XX  
 XX 18-JUN-1999; 99US-0140264.  
 PR 14-SEP-1999; 99US-0153872.  
 PR 19-NOV-1999; 99US-0166573.  
 XX  
 XX (CVTH-) CV THERAPEUTICS INC.  
 XX  
 XX Lawn RM, Wade D, Garvin M;  
 XX WPI; 2001-137812/14.  
 XX  
 XX Adenosine triphosphate (ATP) binding cassette (ABC) polynucleotide,  
 PT useful for the development of agents for the treatment of heart disease  
 PT and other disorders associated with hypercholesterolemia and  
 PT atherosclerosis -  
 XX  
 XX Disclosure; Page 148-154; 215pp; English.  
 XX  
 XX The present sequence encodes a human adenosine triphosphate (ATP)  
 CC binding cassette protein (ABC) 1 polypeptide, and is isolated from  
 CC a Tangier disease patient. ABC1 resides in cell membranes and utilises  
 CC ATP hydrolysis to transport a wide variety of substrates across the  
 CC plasma membrane. ABC1 is a pivotal protein in the apolipoprotein-mediated  
 CC mobilisation of intracellular cholesterol stores. ABC1 is defective in  
 CC Tangier disease, a genetic disorder characterised by abnormal  
 CC HDL-cholesterol metabolism. The ABC1 gene is localised to chromosome  
 CC 9q22-q31. The ABC1 genes and proteins are useful for developing  
 CC pharmaceutical agents for the treatment of heart disease and other  
 CC disorders associated with hypercholesterolemia and atherosclerosis. The  
 CC genes are useful for developing screening assays to screen for compounds  
 CC that regulate the expression of genes associated with cholesterol  
 CC transport. The genes and proteins are also useful for are also useful  
 CC as diagnostic indicators of cardiovascular disease and other disorders  
 CC associated with hypercholesterolemia.  
 XX  
 XX Sequence 10474 BP; 2906 A; 2305 C; 2416 G; 2843 T; 4 other;  
 SQ

Query Match 5.8%; Score 188; DB 22; Length 10474;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-75;  
 Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2927 ccgcagagccgagccgacctctcccgagctgcgcagcagcagcggagctccg 2986  
 Db 42 ccgcagagccgagccgacctctcccgagctgcgcagcagcagcggagctccg 101  
 QY 2987 cgcacacacagcgggtctcagggcgcttgcctctgtttttcccggtctgttt 3046  
 Db 102 cgcacacacagcgggtctcagggcgcttgcctctgtttttcccggtctgttt 161

QY 3047 tctcccttcccgagagcgttgcacgggttagagaaagagcagcaacacaaaagtg 3106  
 Db 162 tctcccttcccgagagcgttgcacgggttagagaaagagcagcaacacaaaagtg 221  
 QY 3107 gaaacacag 3114  
 Db 222 gaaacacag 229  
 RESULT 18  
 AAF24686  
 ID AAF24686 standard; DNA; 10474 BP.  
 XX  
 XX AAF24686;  
 XX 20-APR-2001 (first entry)  
 XX Nucleotide sequence of ABC1 polypeptide from Tangier disease patient.  
 DE Human; adenosine triphosphate binding cassette protein 1; ABC1;  
 XX apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;  
 KW chromosome 9q22-q31; heart disease; hypercholesterolemia;  
 KW atherosclerosis; cholesterol transport; ss.  
 XX Homo sapiens.  
 OS  
 XX Key Location/Qualifiers  
 FH 323..7108  
 FT /\*tag= a  
 FT /product= "defective ABC1 polypeptide"  
 XX  
 XX WO200078972-A2.  
 XX  
 XX 28-DEC-2000.  
 XX  
 XX 16-JUN-2000; 2000WO-US16765.  
 XX  
 XX 18-JUN-1999; 99US-0140264.  
 PR 14-SEP-1999; 99US-0153872.  
 PR 19-NOV-1999; 99US-0166573.  
 XX  
 XX (CVTH-) CV THERAPEUTICS INC.  
 XX  
 XX Lawn RM, Wade D, Garvin M;  
 XX WPI; 2001-137812/14.  
 XX  
 XX Adenosine triphosphate (ATP) binding cassette (ABC) polynucleotide,  
 PT useful for the development of agents for the treatment of heart disease  
 PT and other disorders associated with hypercholesterolemia and  
 PT atherosclerosis -  
 XX  
 XX Disclosure; Page 170-176; 215pp; English.  
 XX  
 XX The present sequence encodes a human adenosine triphosphate (ATP)  
 CC binding cassette protein (ABC) 1 polypeptide, and is isolated from  
 CC a Tangier disease patient. ABC1 resides in cell membranes and utilises  
 CC ATP hydrolysis to transport a wide variety of substrates across the  
 CC plasma membrane. ABC1 is a pivotal protein in the apolipoprotein-mediated  
 CC mobilisation of intracellular cholesterol stores. ABC1 is defective in  
 CC Tangier disease, a genetic disorder characterised by abnormal  
 CC HDL-cholesterol metabolism. The ABC1 gene is localised to chromosome  
 CC 9q22-q31. The ABC1 genes and proteins are useful for developing  
 CC pharmaceutical agents for the treatment of heart disease and other  
 CC disorders associated with hypercholesterolemia and atherosclerosis. The  
 CC genes are useful for developing screening assays to screen for compounds  
 CC that regulate the expression of genes associated with cholesterol  
 CC transport. The genes and proteins are also useful for are also useful  
 CC as diagnostic indicators of cardiovascular disease and other disorders  
 CC associated with hypercholesterolemia.  
 XX  
 XX Sequence 10474 BP; 2907 A; 2304 C; 2415 G; 2844 T; 4 other;  
 SQ

Query Match 5.8%; Score 188; DB 22; Length 10474;  
Best Local Similarity 100.0%; Pred. No. 1.4e-75;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2927 ccgcagagccgagccaccctctctccgggctgcgcagggcgagggcgaggtccg 2986  
DB 42 ccgcagagccgagccaccctctctccgggctgcgcagggcgagggcgaggtccg 101  
QY 2987 cgcaccaacagagccgttctcaggcgcttgcctgtttttcccggtctgttt 3046  
DB 102 cgcaccaacagagccgttctcaggcgcttgcctgtttttcccggtctgttt 161  
QY 3047 tctcccttctccggaagcgttgcagggttagagaaagagacacacaaaagt 3106  
DB 162 tctcccttctccggaagcgttgcagggttagagaaagagacacacaaaagt 221

QY 3107 gaaaacag 3114  
DB 222 gaaaacag 229

RESULT 19  
AAF24707  
ID AAF24707 standard; DNA; 10474 BP.  
AC AAF24707;  
XX  
DT 20-APR-2001 (first entry)  
DE Nucleotide sequence of ABC1 polypeptide from Tangier disease patient.  
KW Human; adenosine triphosphate binding cassette protein 1; ABC1;  
KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;  
KW chromosome 9q22-9q31; heart disease; hypercholesterolemia;  
KW atherosclerosis; cholesterol transport; ss.  
XX  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT CDS 323..7108  
FT /\*tag= a  
FT /product= "defective ABC1 polypeptide"  
XX  
PN WO200078971-A2.  
XX  
PD 28-DEC-2000.  
XX  
PF 16-JUN-2000; 2000WO-US16591.  
XX  
PR 18-JUN-1999; 99US-0140264.  
PR 14-SEP-1999; 99US-0153872.  
PR 19-NOV-1999; 99US-0166573.  
XX  
PA (CVTH-) CV THERAPEUTICS INC.  
PA (UNIW ) UNIV WASHINGTON.  
XX  
PI Lawn RM, Wade D, Oram JF, Garvin M;  
XX  
WPI; 2001-137811/14.  
DR P-PSDB; AAB31366.  
XX  
PT Adenosine triphosphate (ATP) binding cassette protein (ABC) 1  
PT polynucleotides and polypeptides, useful for treatment of heart disease  
PT and other disorders associated with hypercholesterolemia and  
PT atherosclerosis -  
XX  
PS Claim 27; Page 144-150; 211pp; English.  
XX  
CC The present sequence encodes a human adenosine triphosphate (ATP)  
CC binding cassette protein (ABC) 1 polypeptide, and is isolated from  
CC a Tangier disease patient. ABC1 resides in cell membranes and utilises  
CC ATP hydrolysis to transport a wide variety of substrates across the

CC plasma membrane. ABC1 is a pivotal protein in the apolipoprotein-mediated  
CC mobilisation of intracellular cholesterol stores. ABC1 is defective in  
CC Tangier disease, a genetic disorder characterised by abnormal  
CC HDL-cholesterol metabolism. The ABC1 gene is localised to chromosome  
CC 9q22-9q31. The ABC1 genes and proteins are useful for developing  
CC pharmaceutical agents for the treatment of heart disease and other  
CC disorders associated with hypercholesterolemia and atherosclerosis. The  
CC genes are useful for developing screening assays to screen for compounds  
CC that regulate the expression of genes associated with cholesterol  
CC transport. The genes and proteins are also useful for are also useful  
CC as diagnostic indicators of cardiovascular disease and other disorders  
CC associated with hypercholesterolemia.

XX  
SQ Sequence 10474 BP; 2906 A; 2305 C; 2416 G; 2843 T; 4 other;

Query Match 5.8%; Score 188; DB 22; Length 10474;  
Best Local Similarity 100.0%; Pred. No. 1.4e-75;  
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2927 ccgcagagccgagccaccctctctccgggctgcgcagggcgagggcgaggtccg 2986  
DB 42 ccgcagagccgagccaccctctctccgggctgcgcagggcgagggcgaggtccg 101

QY 2987 cgcaccaacagagccgttctcaggcgcttgcctgtttttcccggtctgttt 3046  
DB 102 cgcaccaacagagccgttctcaggcgcttgcctgtttttcccggtctgttt 161

QY 3047 tctcccttctccggaagcgttgcagggttagagaaagagacacacaaaagt 3106  
DB 162 tctcccttctccggaagcgttgcagggttagagaaagagacacacaaaagt 221

QY 3107 gaaaacag 3114  
DB 222 gaaaacag 229

RESULT 20

AAF24708

ID AAF24708 standard; DNA; 10474 BP.

XX AAF24708;

AC AAF24708;

XX 20-APR-2001 (first entry)

XX Nucleotide sequence of ABC1 polypeptide from Tangier disease patient.  
XX Human; adenosine triphosphate binding cassette protein 1; ABC1;  
KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;  
KW chromosome 9q22-9q31; heart disease; hypercholesterolemia;  
KW atherosclerosis; cholesterol transport; ss.  
XX  
OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 323..7108

XX /\*tag= a

XX /product= "defective ABC1 polypeptide"

XX WO200078971-A2.

XX 28-DEC-2000.

XX 16-JUN-2000; 2000WO-US16591.

XX 18-JUN-1999; 99US-0140264.

XX 14-SEP-1999; 99US-0153872.

XX 19-NOV-1999; 99US-0166573.

XX (CVTH-) CV THERAPEUTICS INC.

XX (UNIW ) UNIV WASHINGTON.

XX Lawn RM, Wade D, Oram JF, Garvin M;

XX	WPI: 2001-137811/14.
DR	P-PSDB; AAB31367.
XX	
PT	Adenosine triphosphate (ATP) binding cassette protein (ABC) 1
PT	polynucleotides and polypeptides, useful for treatment of heart disease
PT	and other disorders associated with hypercholesterolemia and
PT	atherosclerosis -
XX	
PS	Claim 30; Page 165-172; 21pp; English.
XX	
CC	The present sequence encodes a human adenosine triphosphate (ATP)
CC	binding cassette protein (ABC) 1 polypeptide, and is isolated from
CC	a Tangier disease patient. ABC1 resides in cell membranes and utilises
CC	ATP hydrolysis to transport a wide variety of substrates across the
CC	plasma membrane. ABC1 is a pivotal protein in the apolipoprotein-mediated
CC	mobilisation of intracellular cholesterol stores. ABC1 is defective in
CC	Tangier disease, a genetic disorder characterised by abnormal
CC	HDL-cholesterol metabolism. The ABC1 gene is localised to chromosome
CC	9q22-9q31. The ABC1 genes and proteins are useful for developing
CC	pharmaceutical agents for the treatment of heart disease and other
CC	disorders associated with hypercholesterolemia and atherosclerosis. The
CC	genes are useful for developing screening assays to screen for compounds
CC	that regulate the expression of genes associated with cholesterol
CC	transport. The genes and proteins are also useful for are also useful
CC	as diagnostic indicators of cardiovascular disease and other disorders
CC	associated with hypercholesterolemia.
XX	
SQ	Sequence 10474 BP; 2907 A; 2304 C; 2415 G; 2844 T; 4 other;

Query Match	5.8%;	Score 188;	DB 22;	Length 10474;
Best Local Similarity	100.0%;	Pred. No. 1.4e-75;		
Matches 188;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 2927	cgcgagagccgagccgacccctctctccgggctcgcgagggcgagggcgagagctccg	2986		
Db 42	cgcgagagccgagccgacccctctctccgggctcgcgagggcgagggcgagagctccg	101		
QY 2987	cgacacacagagccggttctcagggcgcttgcctctgtttttcccggtctctgttt	3046		
Db 102	cgacacacagagcggtctcagggcgcttgcctctgtttttcccggtctctgttt	161		
QY 3047	tctcccttctccgaagcttgcgaagggtagagaaagagacgcaaacacaaagtgc	3106		
Db 162	tctcccttctccgaagcttgcgaagggtagagaaagagacgcaaacacaaagtgc	221		
QY 3107	gaaaacag 3114			
Db 222	gaaaacag 229			

RESULT	21
AAS04035	
ID	AAS04035 standard; cDNA; 446 BP.
XX	
XX	
AAS04035;	
12-SEP-2001	(first entry)
XX	
XX	
Partial human ABC1	cDNA sequence.
DE	
XX	
Human; ABC1 gene;	atherosclerosis; reverse transport; cholesterol; cardiovascular; neurological; Tangier disease;; LCAT deficiency;; lecithin-cholesterol acetyltransferase; malaria; diabetes; ss.
KW	
KW	
KW	
OS	
Homo sapiens.	
XX	
Key	Location/Qualifiers
CDS	185..438
FT	/tag= a
FT	/product= "Human ABC1 protein, amino acids 1 to 60"
FT	
XX	

PN	WO200130848-A2.
XX	
PD	03-MAY-2001.
XX	
PF	26-OCT-2000; 2000WO-EP10885.
XX	
PR	26-OCT-1999; 99EP-0402668.
PR	01-MAR-2000; 2000US-0186260.
XX	
XX	(AVET ) AVENTIS PHARMA SA.
PA	
XX	
PI	Denefle P, Rosier-Montus M, Arnould-Reguigne I, Prades C, Naudin L;
PI	Lemoinc C, Duverger N, Jaye M, Searfoss GH, Remaiey A, Brewer HB;
PI	Dean M;
XX	
XX	WPI: 2001-316327/33.
DR	P-PSDB; RAU02176.
XX	
PT	New human ABC1 nucleic acids and polypeptides for treating
PT	atherosclerosis, malaria and diabetes -
XX	
PS	Example 2; Page 167; 368pp; English.
XX	
CC	The sequence represents the partial coding sequence of human ABC1,
CC	which encodes amino acids 1-60 of the human ABC1 protein. The nucleic
CC	acid sequence, primers and probes derived from the ABC1 sequence, and
CC	polypeptides and vectors are useful for the prevention of
CC	atherosclerosis, in a subject affected by a dysfunction in the reverse
CC	transport of cholesterol. The polypeptide encoded by the ABC1 gene is
CC	useful for screening for an active ingredient for the prevention or
CC	treatment of a disease resulting from dysfunction in the reverse
CC	transport of cholesterol. The nucleic acids and polypeptides are also
CC	useful for treating and preventing cardiovascular and neurological
CC	pathologies, and other diseases e.g. Tangier disease, lecithin-
CC	cholesterol (LCAT) deficiency, malaria and diabetes.
XX	
SQ	Sequence 446 BP; 96 A; 123 C; 112 G; 115 T; 0 other;

	Query Match	2.8%	Score 92;	DB 22;	Length 446;
	Best Local Similarity	100.0%;	Pred. No. 8.8e-32;		
	Matches 92;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	3023	ctgtttttccccgggtctgttttctccctcttcgcgaagcgcttgccaaggggtagga	3082		
Db	1	ctgtttttccccgggtctgttttctccctcttcgcgaagcgcttgccaaggggtagga	60		
Qy	3093	gaaagagacgcaaacacaaaagtggaaaaacag	3114		
Db	61	gaaagagacgcaaacacaaaagtggaaaaacag	92		

RESULT	22
AAS06120	
ID	AAS06120 standard; cDNA; 9741 BP.
XX	XX
XX	AAS06120;
XX	
DT	12-SEP-2001 (first entry)
XX	
DE	Human ABC1 DNA sequence #1.
XX	
KW	Human; ABC1 gene; atherosclerosis; reverse transport; cholesterol;
KW	cardiovascular; neurological; tangier disease; LCAT deficiency;
KW	lecithin-cholesterol acetyltransferase; malaria; diabetes; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Location/Qualifiers
CDS	185..6967
FT	/tag= a
FT	/product= "Human ABC1 protein"
FT	
YY	



PN WO200130848-A2.  
 XX 03-MAY-2001.  
 XX 26-OCT-2000; 2000WO-EP10886.  
 XX 26-OCT-1999; 99EP-0402668.  
 PR 01-MAR-2000; 2000US-0186260.  
 XX (AVET ) AVENTIS PHARMA SA.  
 XX Deneffe P, Rosier-Montus M, Arnould-Reguigne I, Prades C, Naudin L;  
 PI Lemoine C, Duverger N, Jaye M, Searfoss GH, Remaley A, Brewer HB;  
 PI Dean M;  
 XX WPI: 2001-316327/33.  
 DR P-PSDB; RAU02176.  
 XX New human ABC1 nucleic acids and polypeptides for treating  
 PT atherosclerosis, malaria and diabetes -  
 PT Claim 1; Page 204-208; 368pp; English.  
 XX The sequence represents the coding sequence #1 of human ABC1. The  
 CC nucleic acid sequence, primers and probes derived from the ABC1 sequence,  
 CC and polypeptides and vectors are useful for the prevention of  
 CC atherosclerosis, in a subject affected by a dysfunction in the reverse  
 CC transport of cholesterol. The polypeptide encoded by the ABC1 gene is  
 CC useful for screening for an active ingredient for the prevention or  
 CC treatment of a disease resulting from dysfunction in the reverse  
 CC transport of cholesterol. The nucleic acids and polypeptides are also  
 CC useful for treating and preventing cardiovascular and neurological  
 CC pathologies, and other diseases e.g. Tangier disease, lecithin-  
 CC cholesterol (LCAT) deficiency, malaria and diabetes.  
 XX Sequence 9741 BP; 2650 A; 2180 C; 2290 G; 2620 T; 1 other;  
 SQ

Query Match 2.8%; Score 92; DB 22; Length 9741;  
 Best Local Similarity 100.0%; Pred. No. 8 6e-32;  
 Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 3023 cttgtttttcccggttctgttttctcccttctccggaaggtgttcaagggttagga 3082  
 Db 1 cttgtttttcccggttctgttttctcccttctccggaaggtgttcaagggttagga 60  
 Qy 3083 gaaagagacgcaaacacacaaagtggaaacag 3114  
 Db 61 gaaagagacgcaaacacacaaagtggaaacag 92

RESULT 23  
 AAV57926  
 ID AAV57926 standard; DNA; 235033 BP.  
 XX  
 AC AAV57926;  
 XX  
 DT 23-DEC-1998 (first entry)  
 XX  
 DE Hereditary haemochromatosis subregion from an unaffected individual.  
 XX  
 KW Bovine butyrophilin; BT; human hereditary haemochromatosis; HFE;  
 KW diagnosis; iron metabolism; NPT3; NPT4; RoRet; BTF1; BTF2; BTF3;  
 KW BTF4; BTF5; milk protein; lupus; Sjogren's syndrome; hypophosphatemia;  
 KW type 1 sodium transport gene; ss.  
 XX Homo sapiens.  
 OS  
 XX WO9814466-A1.  
 PN  
 XX 09-APR-1998.  
 PD  
 XX 30-SEP-1997; 97WO-US17658.  
 PF

XX 07-MAY-1997; 97US-0852495.  
 PR 01-OCT-1996; 96US-0724394.  
 XX (PROG-) PROCENTIOR INC.  
 XX Feder JN, Kronmal GS, Lauer PM, Ruddy DA, Thomas WJ;  
 PI Tsuchihashi Z, Wolff RK;  
 XX WPI: 1998-240014/21.  
 XX Hereditary haemochromatosis gene products - used to develop products  
 PT for the diagnosis and treatment of hereditary disorders in iron  
 PT metabolism  
 XX Example 2; Fig 8; 209pp; English.  
 XX The present invention describes hereditary haemochromatosis gene  
 CC products from the human haemochromatosis gene. The present sequence  
 CC represents a hereditary haemochromatosis subregion from an individual  
 CC unaffected by hereditary haemochromatosis (HH). Also described is a  
 CC method to determine the presence or absence of the common hereditary  
 CC haemochromatosis (HFE) gene mutation in an individual comprising:  
 CC (a) providing DNA or RNA from the individual; and (b) assessing the  
 CC DNA or RNA for the presence or absence of a haplotype or genotype where  
 CC the presence or absence of the haplotype genotype indicates the likely  
 CC presence of the HFE gene mutation in the genome of the individual. The  
 CC HFE gene sequences from the present invention can be used to develop  
 CC products for use in the diagnosis and treatment of HFE. The present  
 CC invention also describes BTF genes, which are homologues of the milk  
 CC protein butyrophilin (BT), and can be used in the production of agonists  
 CC and antagonists of BT function. Also described are: (1) a RoRet gene  
 CC which can be used to develop products for the study, diagnosis and  
 CC treatment of lupus and Sjogren's syndrome; and (2) NPT3 and NPT4 genes  
 CC which are homologues of a type 1 sodium transport gene, and can  
 CC similarly be used for hypophosphatemia.  
 XX  
 SQ Sequence 235033 BP; 68786 A; 48466 C; 49441 G; 68340 T; 0 other;  
 Query Match 1.7%; Score 55; DB 19; Length 235033;  
 Best Local Similarity 100.0%; Pred. No. 6.4e-15;  
 Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1190 cccacattggtctcccaagtgctggattacaggcatgagccactgcgccagc 1244  
 Db 225009 cccacattggtctcccaagtgctggattacaggcatgagccactgcgccagc 225063

RESULT 24  
 AAV57903  
 ID AAV57903 standard; DNA; 237326 BP.  
 XX  
 AC AAV57903;  
 XX  
 DT 21-DEC-1998 (first entry)  
 XX  
 DE Hereditary haemochromatosis subregion from an HH affected individual.  
 XX  
 KW Bovine butyrophilin; BT; human hereditary haemochromatosis; HFE;  
 KW diagnosis; iron metabolism; NPT3; NPT4; RoRet; BTF1; BTF2; BTF3;  
 KW BTF4; BTF5; milk protein; lupus; Sjogren's syndrome; hypophosphatemia;  
 KW type 1 sodium transport gene; ss.  
 XX Homo sapiens.  
 OS  
 XX WO9814466-A1.  
 PN  
 XX 09-APR-1998.  
 PD  
 XX 30-SEP-1997; 97WO-US17658.  
 PF  
 XX 07-MAY-1997; 97US-0852495.  
 PR

PR 01-OCT-1996; 96US-0724394.  
 XX (PROG-) PROGENITOR INC.  
 XX Feder JN, Kronmal GS, Lauer PM, Ruddy DA, Thomas WJ;  
 PI Tsuchihashi Z, Wolff RK;  
 XX WPI; 1998-240014/21.  
 XX Hereditary haemochromatosis gene products - used to develop products  
 PT for the diagnosis and treatment of hereditary disorders in iron  
 PT metabolism  
 XX Claim 1; Fig 9; 209pp; English.  
 PS The present invention describes hereditary haemochromatosis gene  
 CC products from the human haemochromatosis gene. The present sequence  
 CC represents a hereditary haemochromatosis subregion from an hereditary  
 CC haemochromatosis (HH) affected individual. Also described is a  
 CC method to determine the presence or absence of the common hereditary  
 CC haemochromatosis (HFE) gene mutation in an individual comprising:  
 CC (a) providing DNA or RNA from the individual; and (b) assessing the  
 CC DNA or RNA for the presence or absence of a haplotype or genotype where  
 CC the presence or absence of the haplotype genotype indicates the likely  
 CC presence of the HFE gene mutation in the genome of the individual. The  
 CC HFE gene sequences from the present invention can be used to develop  
 CC products for use in the diagnosis and treatment of HFE. The present  
 CC invention also describes BTF genes, which are homologues of the milk  
 CC protein butyrophilin (BT), and can be used in the production of agonists  
 CC and antagonists of BT function. Also described are: (1) a Roret gene  
 CC which can be used to develop products for the study, diagnosis and  
 CC treatment of lupus and Sjogren's syndrome; and (2) NPT3 and NPT4 genes  
 CC which are homologues of a type 1 sodium transport gene, and can  
 CC similarly be used for hypophosphatemia.  
 XX  
 SQ Sequence 237326 BP; 69596 A; 48904 C; 48217 G; 70609 T; 0 other;

Query Match 1.7%; Score 55; DB 19; Length 237326;  
 Best Local Similarity 100.0%; Pred. No. 6.4e-15;  
 Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1190 ccacattggcctcccaaaagtgtggattacaggaatgagccactgcgccagc 1244  
 |||||  
 Db 227302 ccacattggcctcccaaaagtgtggattacaggaatgagccactgcgccagc 227356  
 |||||

RESULT 25  
 AAI64613  
 ID AAI64613 standard; cDNA; 1316 BP.  
 XX  
 AC AAI64613;  
 XX  
 DT 04-DEC-2001 (first entry)  
 XX  
 DE Human ribosomal S11 protein 12 encoding cDNA.  
 XX  
 KW Human; ribosomal S11 protein 12; cytosolic; virucidal;  
 KW immunomodulatory; antiinflammatory; haemostatic; malignant tumour;  
 KW human immunodeficiency virus; HIV; infection; immunological disease;  
 KW gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH 643..969  
 CDS /\*tag= a  
 FT /product= "ribosomal S11 protein 12"  
 FT /note= "claimed in claim 6"  
 XX  
 XX W0200172801-A1.  
 XX  
 PD 04-OCT-2001.

XX 26-MAR-2001; 2001WO-CN00438.  
 XX  
 XX 27-MAR-2000; 2000CN-0115182.  
 XX  
 XX (SHAN-) SHANGHAI BOWINDOW GENE DEV INC.  
 XX  
 XX Mao Y, Xie Y;  
 XX  
 XX WPI; 2001-597104/67.  
 DR P-PSDB; AAG78169.  
 DR  
 XX New human ribosomal S11 protein 12 and encoded polynucleotide,  
 PT applicable in diagnosis and treatment of malignant tumour, haemopathy,  
 PT human immunodeficiency virus infection, immunological diseases and  
 PT inflammation  
 PT  
 XX Claim 6; Page 29-30; 34pp; Chinese.  
 PS  
 XX The invention relates to the human ribosomal S11 protein 12 with  
 CC cytosolic, virucidal, immunomodulatory, antiinflammatory and  
 CC haemostatic activity. The protein and encoding polynucleotide are used  
 CC in diagnosis and treatment of malignant tumour, haemopathy, human  
 CC immunodeficiency virus (HIV) infection, immunological diseases and  
 CC various inflammations. The polynucleotide is useful in gene therapy.  
 XX  
 SQ Sequence 1316 BP; 338 A; 306 C; 297 G; 375 T; 0 other;

Query Match 1.6%; Score 52; DB 22; Length 1316;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-13;  
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1193 accttggcctcccaaaagtgtggattacaggaatgagccactgcgccagc 1244  
 |||||  
 Db 977 accttggcctcccaaaagtgtggattacaggaatgagccactgcgccagc 1028  
 |||||

RESULT 26  
 AAS31466/c  
 ID AAS31466 standard; DNA; 5351 BP.  
 XX  
 AC AAS31466;  
 XX  
 DT 04-DEC-2001 (first entry)  
 XX  
 DE Human DNA for a novel extracellular matrix protein, Seq ID No 545.  
 XX  
 KW Human; secreted extracellular matrix protein; ds; immunomodulatory;  
 KW Anti-HIV; antianemic; antirheumatic; antisclerotic; cardiac; vascular;  
 KW cerebroprotective; thrombolytic; antimicrobial; ophthalmic; cytostatic;  
 KW antialzheimers; immune/autoimmune disease; HIV infection; anaemia;  
 KW human immunodeficiency virus; rheumatoid arthritis; multiple sclerosis;  
 KW cancers; hyperproliferative disorder; breast neoplasm; melanoma;  
 KW Sezary syndrome; Gaucher's disease; neurological diseases;  
 KW Alzheimer's disease; Parkinson's disease; cardiovascular disorder;  
 KW cardiac arrest; tachycardia; angina; infection; corneal infections;  
 KW wound healing; immunogen; gene therapy; antisense; food additive.  
 XX  
 OS Homo sapiens.  
 XX  
 XX W0200155368-A1.  
 PN  
 PD 02-AUG-2001.  
 XX  
 XX 17-JAN-2001; 2001WO-US01348.  
 XX  
 XX 31-JAN-2000; 2000US-0179065.  
 PR 04-FEB-2000; 2000US-0180628.  
 PR 24-FEB-2000; 2000US-0184664.  
 PR 02-MAR-2000; 2000US-0186350.  
 PR 16-MAR-2000; 2000US-0189874.  
 PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.  
 PR 19-MAY-2000; 2000US-0205515.  
 PR 07-JUN-2000; 2000US-0209467.  
 PR 28-JUN-2000; 2000US-0214886.  
 PR 30-JUN-2000; 2000US-0215135.  
 PR 07-JUL-2000; 2000US-0216647.  
 PR 07-JUL-2000; 2000US-0216880.  
 PR 11-JUL-2000; 2000US-0217487.  
 PR 11-JUL-2000; 2000US-0217496.  
 PR 14-JUL-2000; 2000US-0218290.  
 PR 26-JUL-2000; 2000US-0220963.  
 PR 26-JUL-2000; 2000US-0220964.  
 PR 14-AUG-2000; 2000US-0224518.  
 PR 14-AUG-2000; 2000US-0224519.  
 PR 14-AUG-2000; 2000US-0225213.  
 PR 14-AUG-2000; 2000US-0225214.  
 PR 14-AUG-2000; 2000US-0225266.  
 PR 14-AUG-2000; 2000US-0225267.  
 PR 14-AUG-2000; 2000US-0225268.  
 PR 14-AUG-2000; 2000US-0225270.  
 PR 14-AUG-2000; 2000US-0225447.  
 PR 14-AUG-2000; 2000US-0225757.  
 PR 14-AUG-2000; 2000US-0225758.  
 PR 14-AUG-2000; 2000US-0225759.  
 PR 18-AUG-2000; 2000US-0226279.  
 PR 22-AUG-2000; 2000US-0226681.  
 PR 22-AUG-2000; 2000US-0226868.  
 PR 22-AUG-2000; 2000US-0227182.  
 PR 23-AUG-2000; 2000US-0227009.  
 PR 30-AUG-2000; 2000US-0228924.  
 PR 01-SEP-2000; 2000US-0229287.  
 PR 01-SEP-2000; 2000US-0229343.  
 PR 01-SEP-2000; 2000US-0229344.  
 PR 01-SEP-2000; 2000US-0229345.  
 PR 05-SEP-2000; 2000US-0229509.  
 PR 05-SEP-2000; 2000US-0229513.  
 PR 06-SEP-2000; 2000US-0230437.  
 PR 06-SEP-2000; 2000US-0230438.  
 PR 08-SEP-2000; 2000US-0231242.  
 PR 08-SEP-2000; 2000US-0231243.  
 PR 08-SEP-2000; 2000US-0231244.  
 PR 08-SEP-2000; 2000US-0231413.  
 PR 08-SEP-2000; 2000US-0231414.  
 PR 08-SEP-2000; 2000US-0232080.  
 PR 08-SEP-2000; 2000US-0232081.  
 PR 12-SEP-2000; 2000US-0232168.  
 PR 14-SEP-2000; 2000US-0232397.  
 PR 14-SEP-2000; 2000US-0232398.  
 PR 14-SEP-2000; 2000US-0232399.  
 PR 14-SEP-2000; 2000US-0232399.  
 PR 14-SEP-2000; 2000US-0232401.  
 PR 14-SEP-2000; 2000US-0233064.  
 PR 14-SEP-2000; 2000US-0233065.  
 PR 21-SEP-2000; 2000US-0234223.  
 PR 21-SEP-2000; 2000US-0234274.  
 PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.

PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 01-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-465572/50.

Nucleic acid molecules encoding human secreted extracellular matrix proteins, used in preventing, treating or ameliorating a disorder, e.g. Alzheimer's and Parkinson's diseases and cancers -

Claim 1; SEQ ID No 545; 577pp; English.

The invention relates to isolated nucleic acid molecules encoding novel human secreted extracellular matrix proteins (SPs). The polynucleotides and proteins are used to prevent, treat a medical

CC condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,  
CC chickens or sheep. For example, disorders associated with decreased  
CC expression of Sps. The SP polynucleotide or a vector expressing them may  
CC be administered to treat diseases by gene therapy. Antisense molecules  
CC may be administered to down regulate expression of Sps by binding with  
CC the cells own genes and preventing their expression. The polynucleotides  
CC may also be used as DNA probes in diagnostic assays. The Sps may also be  
CC used as antigens to produce antibodies and to identify modulators  
CC (agonists and antagonists) of the Sps. The anti-(SP) antibodies and  
CC antagonists may also be used to down regulate expression and activity of  
CC SP and as diagnostic agents for detecting the presence of Sps in samples.  
CC The disorders include for example: immune/autoimmune diseases (e.g. HIV  
CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis  
CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.  
CC melanomas, neoplasms of the breast or liver, Sezary syndrome and  
CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,  
CC Parkinson's disease) cardio-/cerebrovascular disorders (e.g. cardiac  
CC arrest, tachycardia and angina), infections caused by bacteria, viruses  
CC and fungi and ocular disorders (e.g. corneal infections). Other uses  
CC include wound healing, maintenance of organs before transplantation,  
CC support of cell culture of primary tissues, modulation of for example

Query Match 1.68; Score 52; DB 22; Length 5351;

Best Local Similarity 100.0%; Pred. No. 1.5e-13;

Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1193 accttggctcccaagtgctggattacaggcatgagccactgcgcccagc 1244

|||||

Db 989 ACCTTGGCTCCCAAGTGGTGGATTACAGGCATGAGCCACATGCGCCCGAGC 938

RESULT 27

AAS31467/c

ID AAS31467 standard; DNA; 6461 BP.

XX AC AAS31467;

XX DT 04-DEC-2001 (first entry)

XX DE Human DNA for a novel extracellular matrix protein, Seq ID No 546.

KW Human; secreted extracellular matrix protein; ds; immunomodulatory;  
KW Anti-HIV; antianemic; antirheumatic; antisclerotic; cardiac; vascular;  
KW cerebroprotective; thrombolytic; antimicrobial; ophthalmic; cytostatic;  
KW antialzheimers; immune/autoimmune disease; HIV infection; anaemia;  
KW human immunodeficiency virus; rheumatoid arthritis; multiple sclerosis;  
KW cancers; hyperproliferative disorder; breast neoplasm; melanoma;  
KW Sezary syndrome; Gaucher's disease; neurological diseases;  
KW Alzheimer's disease; Parkinson's disease; cardiovascular disorder;  
KW cardiac arrest; tachycardia; angina; infection; corneal infections;  
KW wound healing; immunogen; gene therapy; antisense; food additive.

XX OS Homo sapiens.

XX PN WO200153368-A1.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US01348.

XX PR 31-JAN-2000; 2000US-0179065.

XX PR 04-FEB-2000; 2000US-0180628.

XX PR 24-FEB-2000; 2000US-0184664.

XX PR 02-MAR-2000; 2000US-0186350.

XX PR 16-MAR-2000; 2000US-0189874.

XX PR 17-MAR-2000; 2000US-0190076.

XX PR 18-APR-2000; 2000US-0198123.

XX PR 19-MAY-2000; 2000US-0205515.

XX PR 07-JUN-2000; 2000US-0209467.

XX PR 28-JUN-2000; 2000US-0214886.

XX PR 30-JUN-2000; 2000US-0215135.

XX PR 07-JUL-2000; 2000US-0216647.

XX PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.

PR 20-OCT-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-465572/50.

XX Nucleic acid molecules encoding human secreted extracellular matrix  
 PT proteins, used in preventing, treating or ameliorating a disorder, e.g.  
 PT Alzheimer's and Parkinson's diseases and cancers -

PS Claim 1; SEQ ID No 546; 577pp; English.

XX The invention relates to isolated nucleic acid molecules encoding  
 CC novel human secreted extracellular matrix proteins (SPs). The  
 CC polynucleotides and proteins are used to prevent, treat a medical  
 CC condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,  
 CC chickens or sheep. For example, disorders associated with decreased  
 CC expression of SPs. The SP polynucleotide or a vector expressing them may  
 CC be administered to treat diseases by gene therapy. Antisense molecules  
 CC may be administered to down regulate expression of SPs by binding with  
 CC the cells own genes and preventing their expression. The polynucleotides  
 CC may also be used as DNA probes in diagnostic assays. The SPs may also be

CC used as antigens to produce antibodies and to identify modulators  
 CC (agonists and antagonists) of the SPs. The anti-(SP) antibodies and  
 CC antagonists may also be used to down regulate expression and activity of  
 CC SP and as diagnostic agents for detecting the presence of SPs in samples.  
 CC The disorders include for example: immune/autoimmune diseases (e.g. HIV  
 CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis  
 CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.  
 CC melanomas, neoplasms of the breast or liver, Sezary syndrome and  
 CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,  
 CC Parkinson's disease) cardio-/cerebrovascular disorders (e.g. cardiac  
 CC arrest, tachycardia and angina), infections caused by bacteria, viruses  
 CC and fungi and ocular disorders (e.g. corneal infections). Other uses  
 CC include wound healing, maintenance of organs before transplantation,  
 CC support of cell culture of primary tissues, modulation of for example

Query Match 1.6%; Score 52; DB 22; Length 6461;

Best Local Similarity 100.0%; Pred. No. 1.5e-13;

Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1193 acctggcctcccaagtgctgggattacaggcatgagccactgcgccagc 1244  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 4277 ACCTTGGCCTCCCAAGTGTGGGATTACAGGATGAGCCACTGCGCCAGC 4226

RESULT 28

AAK65197

ID AAK65197 standard; DNA; 8319 BP.

XX AAK65197;

XX 06-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:20009.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
 KW cytosstatic; gene therapy; vaccine; metastasis; ds.

XX Homo sapiens.

XX WO200157182-A2.

XX 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US01354.

XX 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 26-JUL-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225266.

PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.

PR 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225757.

PR 14-AUG-2000; 2000US-02257558.  
PR 14-AUG-2000; 2000US-02257559.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 03-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 12-SEP-2000; 2000US-0232081.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.

PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251866.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 11-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483426/52.

Nucleic acids encoding human immune/hematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis -

Disclosure; SEQ ID NO 20009; 3071pp + Sequence Listing; English.

AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I) amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate (I) expression. For example, they may be used to treat disorders associated with decreased expression by rectifying mutations or deletions in a patient's genome that affect the activity of (I) by expressing inactive proteins or to supplement the patient's own production of (I). Additionally, (I) polynucleotides may be used to produce the secreted (I), by inserting the nucleic acids into a host cell and culturing the cell to express the protein. (I) proteins and polynucleotides may be used to prevent, diagnose and treat immune/haematopoietic-related diseases, especially cancers and cancer metastases of haematopoietic-derived cells. AAK64703 to AAK87694 represent human immune/haematopoietic antigen genomic sequences from the present invention. AAK54942 to AAK54950 and AAK82169 represent sequences used in the exemplification of the present invention.

Sequence 8319 BP; 2521 A; 1610 C; 1560 G; 2628 T; 0 other;

Query Match 1.6%; Score 52; DB 22; Length 8319;  
Best Local Similarity 100.0%; Pred. No. 1.5e-13;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;







PS Disclosure; SEQ ID NO 10357; 1701pp + Sequence Listing; English.

XX The invention relates to novel genes (ABA11004-ABA21534) and proteins (AB114678-AB18001) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.

CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 29329 BP; 5864 A; 6756 C; 7225 G; 8484 T; 0 other;

SQ

Query Match 1.6%; Score 52; DB 22; Length 29329;  
Best Local Similarity 100.0%; Pred. No. 1.5e-13;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1187 ctgcccacattgctcccaagtgctgggattacagcagcactgcg 1238  
|||||  
DB 22210 CTGCCACCTTGCTCCCAAGTGTGGATTACAGCATGAGCCACTGCG 22159

RESULT 31  
ABA20511/c  
ID ABA20511 standard; DNA; 29329 BP.

XX  
AC ABA20511;  
XX  
DT 23-JAN-2002 (first entry)  
XX  
DE Human nervous system related polynucleotide SEQ ID NO 12842.  
XX  
KW Human; nootropic; neuroprotective; cytostatic; dermatological; virucide; immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnary; antiparkinsonian; antisickling; antianemic; antiarthritic; cancer; antirheumatic; hepatotropic; cerebroprotective; antiinflammatory; antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal; antiparasitic; cardiac; immune disorder; cardiovascular disorder; neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.

OS Homo sapiens.  
XX  
XX WO200159063-A2.  
XX  
PD 16-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US01334.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.

PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 20-OCT-2000; 2000US-0242221.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250391.  
PR 01-DEC-2000; 2000US-0251160.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-541565/60.  
XX  
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating nervous system  
PT cancers and metastases -  
XX

PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225271.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226686.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 12-SEP-2000; 2000US-0232081.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 13-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 20-OCT-2000; 2000US-0242221.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250391.  
PR 01-DEC-2000; 2000US-0251160.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 11-DEC-2000; 2000US-0251990.  
PR 05-JAN-2001; 2000US-0254097.  
PR 05-JAN-2001; 2000US-0259678.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-541565/60.  
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,  
XX useful for preventing, diagnosing and/or treating nervous system  
XX cancers and metastases -  
XX Disclosure; SEQ ID NO 12842; 1701pp + Sequence Listing; English.  
PS The invention relates to novel genes (ABAI004-ABA21534) and proteins  
XX (ABBI4678-ABBI8001) useful for preventing, treating or ameliorating  
CC medical conditions e.g. by protein or gene therapy. The genes are  
CC isolated from a range of human tissues disclosed in the specification.  
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful  
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone  
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative

CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;  
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
 CC and parasitic infections.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences).  
 XX  
 SQ Sequence 29329 BP; 6864 A; 6756 C; 7225 G; 8484 T; 0 other;

Query Match 1.6%; Score 52; DB 22; Length 29329;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-13;  
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1187 ctgccaccttgccctcccaagctggtgattacagggcatgagccactgcy 1238  
 |||||||  
 Db 22210 CTGCCACCTTGCCCTCCCAAGTCTGGATTACAGGCATGAGCCACTGCG 22159

RESULT 32  
 AAK70791/C  
 ID AAK70791 standard; DNA; 29329 BP.  
 XX  
 AC AAK70791;  
 XX  
 XX 06-NOV-2001 (first entry)  
 XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:25603.  
 XX Human; Immune; haematopoietic; immune/haematopoietic antigen; cancer;  
 KW Cytostatic; gene therapy; vaccine; metastasis; ds.  
 XX Homo sapiens.  
 OS  
 XX WO200157182-A2.  
 PN  
 XX 09-AUG-2001.  
 PD  
 XX 17-JAN-2001; 2001WO-US01354.  
 XX

XX 31-JAN-2000; 2000US-0179065.  
 PR 04-FEB-2000; 2000US-0180628.  
 PR 24-FEB-2000; 2000US-0184664.  
 PR 02-MAR-2000; 2000US-0186350.  
 PR 16-MAR-2000; 2000US-0189874.  
 PR 17-MAR-2000; 2000US-0190076.  
 PR 18-APR-2000; 2000US-0198123.  
 PR 19-MAY-2000; 2000US-0205515.  
 PR 07-JUN-2000; 2000US-0209467.  
 PR 28-JUN-2000; 2000US-0214886.  
 PR 30-JUN-2000; 2000US-0215135.  
 PR 07-JUL-2000; 2000US-0216647.  
 PR 07-JUL-2000; 2000US-0216880.  
 PR 11-JUL-2000; 2000US-0217487.  
 PR 11-JUL-2000; 2000US-0217496.  
 PR 14-JUL-2000; 2000US-0218290.  
 PR 26-JUL-2000; 2000US-0220963.  
 PR 26-JUL-2000; 2000US-0220964.  
 PR 14-AUG-2000; 2000US-0224518.  
 PR 14-AUG-2000; 2000US-0224519.  
 PR 14-AUG-2000; 2000US-0225213.  
 PR 14-AUG-2000; 2000US-0225214.  
 PR 14-AUG-2000; 2000US-0225266.  
 PR 14-AUG-2000; 2000US-0225267.  
 PR 14-AUG-2000; 2000US-0225268.  
 PR 14-AUG-2000; 2000US-0225270.  
 PR 14-AUG-2000; 2000US-0225447.  
 PR 14-AUG-2000; 2000US-0225757.  
 PR 14-AUG-2000; 2000US-0225758.  
 PR 14-AUG-2000; 2000US-0225759.  
 PR 18-AUG-2000; 2000US-0226279.  
 PR 22-AUG-2000; 2000US-0226681.  
 PR 22-AUG-2000; 2000US-0226868.  
 PR 22-AUG-2000; 2000US-0227182.  
 PR 23-AUG-2000; 2000US-0227009.  
 PR 30-AUG-2000; 2000US-0228924.  
 PR 01-SEP-2000; 2000US-0229287.  
 PR 01-SEP-2000; 2000US-0229343.  
 PR 01-SEP-2000; 2000US-0229344.  
 PR 01-SEP-2000; 2000US-0229345.  
 PR 05-SEP-2000; 2000US-0229509.  
 PR 05-SEP-2000; 2000US-0229513.  
 PR 06-SEP-2000; 2000US-0230437.  
 PR 06-SEP-2000; 2000US-0230438.  
 PR 08-SEP-2000; 2000US-0231242.  
 PR 08-SEP-2000; 2000US-0231243.  
 PR 08-SEP-2000; 2000US-0231244.  
 PR 08-SEP-2000; 2000US-0231413.  
 PR 08-SEP-2000; 2000US-0231414.  
 PR 08-SEP-2000; 2000US-0232080.  
 PR 08-SEP-2000; 2000US-0232081.  
 PR 12-SEP-2000; 2000US-0231968.  
 PR 14-SEP-2000; 2000US-0232397.  
 PR 14-SEP-2000; 2000US-0232398.  
 PR 14-SEP-2000; 2000US-0232399.  
 PR 14-SEP-2000; 2000US-0232400.  
 PR 14-SEP-2000; 2000US-0232401.  
 PR 14-SEP-2000; 2000US-0233063.  
 PR 14-SEP-2000; 2000US-0233064.  
 PR 14-SEP-2000; 2000US-0233065.  
 PR 21-SEP-2000; 2000US-0234223.  
 PR 21-SEP-2000; 2000US-0234274.  
 PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 13-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 01-NOV-2000; 2000US-0241826.  
 PR 08-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 17-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.







XX PS Claim 1; SEQ ID NO 8018; 1399pp + Sequence Listing; English.  
XX CC The invention relates to human polynucleotides (AAI79941-AAI93841) and  
CC the encoded proteins (AAO0010-AAO13910) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX SQ Sequence 425 BP; 141 A; 83 C; 84 G; 116 T; 1 other;  
  
Query Match 1.6%; Score 51; DB 22; Length 425;  
Best Local Similarity 100.0%; Pred. No. 4.5e-13;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1194 ccttgccctcccaagctgggtacagcagcatgagccactgcccagc 1244  
|||||  
Db 174 CTTGGCCTCCCAAGTCTGGGATTACAGCATGAGCCACTGCGCCAGC 124  
  
RESULT 37  
AAK77921  
ID AAK77921 standard; DNA; 1856 BP.  
XX AC AAK77921;  
XX DT 07-NOV-2001 (first entry)  
XX DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:32733.  
XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytotatic; gene therapy; vaccine; metastasis; ds.  
XX OS Homo sapiens.  
XX PN WO200157182-A2.  
XX PD 09-AUG-2001.  
XX PF 17-JAN-2001; 2001WO-US01354.  
XX PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226688.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 12-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 13-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.

PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251088.  
PR 05-DEC-2000; 2000US-0251988.  
PR 06-DEC-2000; 2000US-0251719.  
PR 08-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-483426/52.

XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and  
PT metastasis -

XX Disclosure; SEQ ID NO 32733; 3071pp + Sequence Listing; English.

XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting  
CC the nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/hematopoietic-related diseases, especially  
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/hematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
CC represent sequences used in the exemplification of the present invention.

XX Sequence 1856 BP; 375 A; 416 C; 358 G; 707 T; 0 other;

Query Match 1.6%; Score 51; DB 22; Length 1856;  
Best Local Similarity 100.0%; Pred. No. 4.4e-13;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1187 ctgcccaccttgccctcccaagtcgtggattacagcagcatgagccactgc 1237  
|||||  
Db 740 ctgcccaccttgccctcccaagtcgtggattacagcagcatgagccactgc 790

#### RESULT 38

AAI64545/c

ID AAI64545 standard; cDNA; 2046 BP.

XX AC AAI64545;

XX DT 23-NOV-2001 (first entry)

XX Human polypeptide-cytochrome b5-13 encoding cDNA.

XX Human; polypeptide-cytochrome b5-13; malignant tumour; haemopathy; HIV;  
XX human immunodeficiency virus; infection; immunological disease;  
XX inflammation; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 1083..1451

XX /\*tag= a

XX /product= "polypeptide-cytochrome b5-13"

XX /note= "claimed in claim 6"

XX CML301705-A.

XX PD 04-JUL-2001.

XX PF 27-DEC-1999; 99CN-0125352.

XX PR 27-DEC-1999; 99CN-0125352.

XX PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.

XX PI Mao Y, Xie Y;

XX WPI; 2001-550466/62.

XX P-PSDB; AAG78149.

XX New polypeptide-cytochrome b 5-13 and its encoding polynucleotide  
PT useful for treating tumour, immunological disease, haemopathy, human  
PT immunodeficiency virus infection and inflammation -

XX Claim 6; Page 23-24 (Disclosure); 30pp; Chinese.

XX The invention relates to the human polypeptide-cytochrome b5-13 and its  
CC encoding polynucleotide. The polypeptide is used to treat various  
CC diseases, such as malignant tumour, haemopathy, human immunodeficiency  
CC virus infection, immunological diseases and various inflammation.

XX Sequence 2046 BP; 608 A; 472 C; 411 G; 555 T; 0 other;

Query Match 1.6%; Score 51; DB 22; Length 2046;  
Best Local Similarity 100.0%; Pred. No. 4.4e-13;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1194 cttggcctcccaagtcgtggattacagcagcatgagccactgcgccagc 1244  
|||||  
Db 1542 CCTGTGGCTCCCAAGTCTGGGATTACAGGCATGAGCCACTGCGCCAGC 1492

#### RESULT 39

AAK77916



ID AAK77916 standard; DNA; 7759 BP.  
XX  
AC AAK77916;  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:32728.  
XX  
DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis; ds.  
XX Homo sapiens.  
XX  
PN WO200157182-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01354.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226686.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249266.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR



PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 01-NOV-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 (HUMA-) HUMAN GENOME SCI INC.  
 Rosen CA, Barash SC, Ruben SM;  
 WPI; 2001-488786/53.

PT New isolated ovarian and/or breast cancer related nucleic acids and  
 PT polypeptides, useful for diagnosing, treating and/or preventing human  
 PT diseases and disorders, particularly ovarian and/or breast cancer -  
 XX Disclosure; SEQ ID NO 708; 577pp + Sequence Listing; English.  
 XX  
 CC The invention relates to novel genes (ABA07454-ABA08224) and proteins  
 CC (ABB10743-ABB10980) useful for preventing, treating or ameliorating  
 CC medical conditions e.g. by protein or gene therapy. The genes are  
 CC isolated from a range of human tissues disclosed in the specification.  
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful  
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone  
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
 CC hemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;  
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
 CC and parasitic infections.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_sequences.  
 XX  
 SQ Sequence 17904 BP; 5485 A; 3834 C; 3913 G; 4672 T; 0 other;  
 Query Match 1.6%; Score 51; DB 22; Length 17904;  
 Best Local Similarity 100.0%; Pred. No. 4.4e-13;  
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1190 cceacttgccctcccaagctggtgattacagcagcactgacccgc 1240  
 Db 4315 CCCACCTTGGCCCTCCCAAGTCTGGGATTACAGCAGCAGCCTCGGCC 4265  
 RESULT 41  
 AAL03730/G  
 ID AAL03730 standard; DNA; 17904 BP.  
 AC AAL03730;  
 XX  
 XX 21-NOV-2001 (first entry)  
 DE Human reproductive system related antigen DNA SEQ ID NO: 6418.  
 KW Human; reproductive system related antigen; reproductive system disorder;  
 KW cancer; gene therapy; ds.  
 XX Homo sapiens.  
 OS  
 XX WO200155320-A2.  
 XX  
 XX 02-AUG-2001.  
 XX  
 XX 17-JAN-2001; 2001WO-US01339.  
 XX  
 XX 31-JAN-2000; 2000US-0179065.  
 PR 04-FEB-2000; 2000US-0180628.  
 PR 24-FEB-2000; 2000US-0184664.  
 PR 02-MAR-2000; 2000US-0186350.  
 PR 16-MAR-2000; 2000US-0189874.  
 PR 17-MAR-2000; 2000US-0190076.  
 PR 18-APR-2000; 2000US-0198123.  
 PR 19-MAY-2000; 2000US-0205515.  
 PR 07-JUN-2000; 2000US-0209467.  
 PR 28-JUN-2000; 2000US-0214896.  
 PR 30-JUN-2000; 2000US-0215135.  
 PR 07-JUL-2000; 2000US-0216647.  
 PR 07-JUL-2000; 2000US-0216880.  
 PR 11-JUL-2000; 2000US-0217487.  
 PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226688.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231444.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.

PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-465570/50.

Isolated nucleic acid molecule encoding a reproductive system antigen -  
is used in preventing, treating or ameliorating a medical condition -

Disclosure; SEQ ID NO 6418; 1297pp + Sequence Listing; English.

The present invention provides the protein and coding sequences of a  
number of human reproductive system related antigens. These can be used  
in the prevention and treatment of reproductive system disorders,  
including cancer. The present sequence is a genomic sequence encoding a  
protein of the invention.

Sequence 17904 BP; 5485 A; 3834 C; 3913 G; 4672 T; 0 other;

Query Match 1.6%; Score 51; DB 22; Length 17904;  
Best Local Similarity 100.0%; Pred. No. 4.4e-13;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1190 ccacattgctcccaaaagtctgggtacacaggtacagccactgccc 1240  
AA  
Db 4315 CCCACTTGGCCCTCCCAAAAGTGGGATTACAGGCATGAGCCACTGGCC 4265  
  
RESULT 42  
AAL36313/c  
ID AAL36313 standard; DNA; 26591 BP.  
XX  
AC AAL36313;  
XX  
DT 08-JAN-2002 (first entry)  
XX  
DE Human musculoskeletal system related polynucleotide SEQ ID NO 2678.  
XX  
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;  
KW vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;  
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; human; secreted protein;  
KW musculoskeletal system; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200155367-A1.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01338.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226691.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231245.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0234984.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.

PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0256719.  
 PR 08-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Rosen CA, Barash SC, Ruben SM;  
 PI  
 XX  
 XX WPI; 2001-451937/48.  
 DR  
 XX

PT Isolated polypeptide for treating, preventing and/or prognosing  
 PT disorders related to the musculoskeletal system including  
 PT musculoskeletal cancers and also for testing and detection e.g.  
 PT diagnosis -  
 PT  
 XX

PS Example 2; SEQ ID NO 2678; 781pp + Sequence Listing; English.  
 XX

CC The invention relates to novel genes (AAL34669-AAL37666) and proteins  
 CC (ABB03087-ABB04109) associated with the musculoskeletal system useful  
 CC for preventing, treating or ameliorating medical conditions e.g. by  
 CC protein or gene therapy. The genes are isolated from a range of human  
 CC tissues disclosed in the specification. The nucleic acids, proteins,  
 CC antibodies and (ant)agonists are useful in the diagnosis, treatment  
 CC and prevention of: (a) cancer, e.g. breast and ovarian cancer and  
 CC other cancers of the adrenal gland, bone, bone marrow, breast,  
 CC gastrointestinal tract, liver, lung, or urogenital; (b) immune  
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;  
 CC (c) cardiovascular disorders such as myocardial ischaemias; (d) wound  
 CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;  
 CC and (f) infectious diseases such as viral, bacterial, fungal and  
 CC parasitic infections.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 26591 BP; 7489 A; 6354 C; 5953 G; 6795 T; 0 other;

Query Match 1.6%; Score 51; DB 22; Length 26591;  
 Best Local Similarity 100.0%; Pred. No. 4.4e-13;  
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1187 ctgccacccttgctcccaagtgctgattacagcagcagcactgc 1237  
 ||||||||||||||||||||||||||||||||||||||||||||  
 Db 11281 CTGCCACCTTGGCTCCCAAGTGTGGGATTACAGCAGCAGCCTGC 11231

RESULT 43  
 AAS34422  
 ID AAS34422 standard; DNA; 32186 BP.  
 XX  
 AC AAS34422;  
 XX  
 DT 17-DEC-2001 (first entry)

XX  
 DE Human DNA for a novel foetal antigen, SEQ ID No 1846.  
 XX  
 XX Human; foetal tissue antigen; ds; antiinflammatory; neuroprotective;  
 KW immunomodulator; cardiovascular; cytostatic; nephrothropic;  
 KW cardiovascular; autoimmune disease; rheumatoid arthritis;  
 KW hyperproliferative disorder; breast neoplasm; cancer;  
 KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;  
 KW cerebral ischaemia; angiogenesis; nervous system disorder;  
 KW Alzheimer's disease; infection; ocular proliferation; food additive.  
 KW wound healing; epithelial cell proliferation; food additive.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200155312-A2.  
 PN  
 XX  
 XX 02-AUG-2001.  
 PD  
 XX  
 XX 17-JAN-2001; 2001WO-US01321.  
 PF  
 XX  
 XX 31-JAN-2000; 2000US-0179065.  
 PR 04-FEB-2000; 2000US-0180628.  
 PR 24-FEB-2000; 2000US-0184664.  
 PR 02-MAR-2000; 2000US-0186350.  
 PR 16-MAR-2000; 2000US-0189874.  
 PR 17-MAR-2000; 2000US-0190076.  
 PR 18-APR-2000; 2000US-0198123.  
 PR 19-MAY-2000; 2000US-0205515.  
 PR 07-JUN-2000; 2000US-0209467.  
 PR 28-JUN-2000; 2000US-0214886.  
 PR 30-JUN-2000; 2000US-0215135.  
 PR 07-JUL-2000; 2000US-0216647.  
 PR 11-JUL-2000; 2000US-0216880.  
 PR 11-JUL-2000; 2000US-0217487.  
 PR 14-JUL-2000; 2000US-0217496.  
 PR 14-JUL-2000; 2000US-0218290.  
 PR 26-JUL-2000; 2000US-0220963.  
 PR 26-JUL-2000; 2000US-0220964.  
 PR 14-AUG-2000; 2000US-0224518.  
 PR 14-AUG-2000; 2000US-0224519.  
 PR 14-AUG-2000; 2000US-0225213.  
 PR 14-AUG-2000; 2000US-0225214.  
 PR 14-AUG-2000; 2000US-0225266.  
 PR 14-AUG-2000; 2000US-0225267.  
 PR 14-AUG-2000; 2000US-0225268.  
 PR 14-AUG-2000; 2000US-0225270.  
 PR 14-AUG-2000; 2000US-0225447.  
 PR 14-AUG-2000; 2000US-0225757.  
 PR 14-AUG-2000; 2000US-0225758.  
 PR 14-AUG-2000; 2000US-0225759.  
 PR 18-AUG-2000; 2000US-0226279.  
 PR 22-AUG-2000; 2000US-0226681.  
 PR 22-AUG-2000; 2000US-0226868.  
 PR 22-AUG-2000; 2000US-0227182.  
 PR 23-AUG-2000; 2000US-0227009.  
 PR 30-AUG-2000; 2000US-0228924.  
 PR 01-SEP-2000; 2000US-0229287.  
 PR 01-SEP-2000; 2000US-0229343.  
 PR 01-SEP-2000; 2000US-0229344.  
 PR 01-SEP-2000; 2000US-0229345.  
 PR 05-SEP-2000; 2000US-0229509.  
 PR 05-SEP-2000; 2000US-0229513.  
 PR 06-SEP-2000; 2000US-0230437.  
 PR 06-SEP-2000; 2000US-0230438.  
 PR 08-SEP-2000; 2000US-0231242.  
 PR 08-SEP-2000; 2000US-0231243.  
 PR 08-SEP-2000; 2000US-0231244.  
 PR 08-SEP-2000; 2000US-0231413.  
 PR 08-SEP-2000; 2000US-0231414.  
 PR 08-SEP-2000; 2000US-0232080.  
 PR 08-SEP-2000; 2000US-0232081.  
 PR 12-SEP-2000; 2000US-0231968.  
 PR 14-SEP-2000; 2000US-0232397.



OS Homo sapiens.  
 PN WO200026382-A1.  
 XX 11-MAY-2000.  
 PD  
 XX  
 PF 28-OCT-1999; 99WO-JP05983.  
 XX  
 PR 30-OCT-1998; 98JP-0310422.  
 XX  
 XX (MEDI-) MEDICAL & BIOLOGICAL LAB CO LTD.  
 XX  
 XX Toji S, Yano M, Tamai K;  
 PI  
 XX WPI; 2000-365627/31.  
 DR  
 XX Thiorodoxin reductase II (an X-linked inhibitor for apoptosis protein  
 PT (XIAP)-binding protein), with thiorodoxin reductase activity, useful  
 PT e.g. for treating apoptosis-related disorders, cancer and inflammation  
 PT  
 XX  
 PS Claim 15; Page 74-135; 139pp; Japanese.  
 XX  
 XX This sequence represents a human thiorodoxin reductase (Txr) encoding  
 CC genomic nucleotide sequence. The invention relates to two Txr proteins,  
 CC which have XIAP (X-linked inhibitor of apoptosis protein)-binding protein  
 CC activity. The human Txr gene is located on chromosome 22q11.2. The  
 CC invention includes antibodies which bind to the proteins, a vector  
 CC containing the Txr encoding nucleotide sequences, and methods for  
 CC producing transformants using the vector. Thiorodoxin reductase has  
 CC cytostatic and anti-inflammatory activity, and is used in the treatment  
 CC of diseases relating to apoptosis particularly due to cancer of viral  
 CC infection. Txr can also be used to treat inflammation and in the  
 CC screening of anti-cancer agents.  
 XX  
 SQ Sequence 66566 BP; 13694 A; 17724 C; 18303 G; 16844 T; 1 other;  
 XX

Query Match 1.6%; Score 51; DB 21; Length 66566;  
 Best Local Similarity 100.0%; Pred. No. 4.3e-13;  
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1194 ccttggcctcccaagctgctggattacagcagcagcactgcgcccagc 1244  
 |||||  
 Db 1791 ccttggcctcccaagctgctggattacagcagcagcactgcgcccagc 1841

RESULT 45  
 AAC89560  
 ID AAC89560 standard; DNA; 122186 BP.  
 XX  
 AC AAC89560;  
 XX  
 XX 08-MAR-2001 (first entry)  
 DT  
 DE Human histone deacetylase HDAC-D coding sequence.  
 XX  
 KW Histone deacetylase; HDAC-1; HDAC-2; HDAC-3; HDAC-4; HDAC-5; HDAC-C;  
 KW HDAC-D; cell cycle; tumorigenesis; cancer; inhibitor; antisense;  
 KW gene therapy; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200071703-A2.  
 XX  
 PD 30-NOV-2000.  
 XX  
 PF 03-MAY-2000; 2000WO-IB01252.  
 XX  
 PR 03-MAY-1999; 99US-0132287.  
 XX  
 PA (METH-) METHYLGENE INC.  
 XX

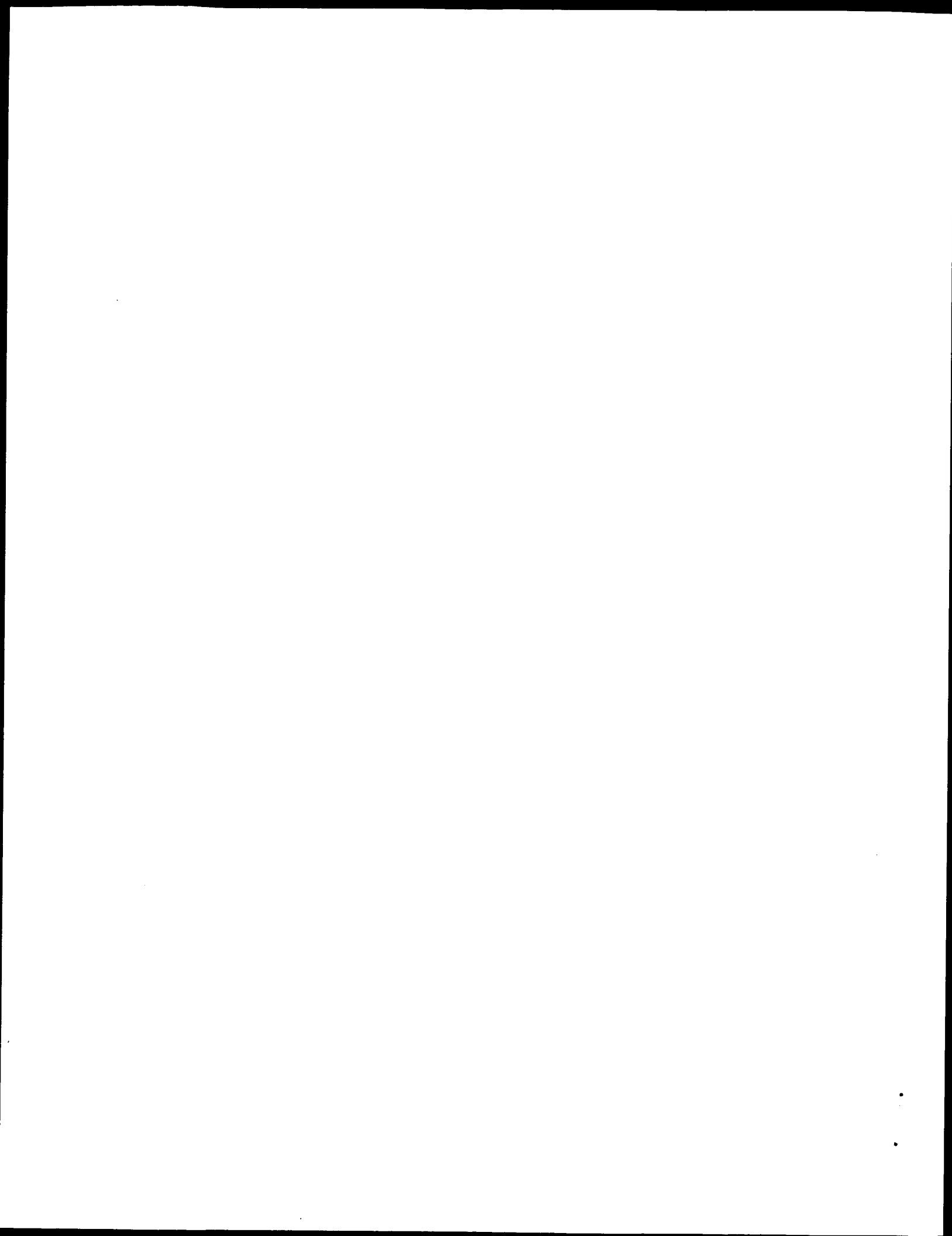
PI Macleod AR, Li Z, Besterman JM;  
 XX  
 DR WPI; 2001-016407/02.  
 XX  
 XX Antisense oligonucleotide that inhibits expression of a histone  
 PT deacetylase, useful for treating and/or alleviating the symptoms of  
 PT neoplasia, or for inhibiting neoplastic cell growth in an animal -  
 XX  
 PS Disclosure; Page 89-125; 125pp; English.  
 XX  
 CC The present invention provides inhibitors of histone deacetylase enzymes  
 CC such as HDAC-1, HDAC-2, HDAC-3, HDAC-4, HDAC-5, HDAC-C and HDAC-D. These  
 CC inhibitors may be antisense strands or they may be compounds identified  
 CC by contacting the enzyme with the compound and measuring the resulting  
 CC enzyme activity. These inhibitors are useful for treating cancers and for  
 CC identifying which histone deacetylase is involved in a neoplasia.  
 XX  
 SQ Sequence 122186 BP; 29016 A; 31077 C; 32425 G; 29668 T; 0 other;  
 XX

Query Match 1.6%; Score 51; DB 22; Length 122186;  
 Best Local Similarity 100.0%; Pred. No. 4.3e-13;  
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1194 ccttggcctcccaagctgctggattacagcagcagcactgcgcccagc 1244  
 |||||  
 Db 65150 ccttggcctcccaagctgctggattacagcagcagcactgcgcccagc 65200

Search completed: September 20, 2002, 06:06:55  
 Job time: 10299 sec







GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 03:09:41 ; Search time 111.56 Seconds  
(without alignments)  
7114.036 Million cell updates/sec

Title: US-09-846-456-1  
Perfect score: 3231  
Sequence: 1 acaggcgtgtgagtg.....gccccacatccccaccatt 3231

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 383533 seqs, 122816752 residues

Word size : 0

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents\_NA:  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	55	1.7	246240	2	US-08-724-394A-20
c 2	55	1.7	246240	2	US-08-724-394A-21
c 3	55	1.7	246240	2	US-08-724-394A-22
c 4	48	1.5	31571	1	US-08-323-443B-1
c 5	48	1.5	53526	3	US-08-658-136-2
c 6	48	1.5	53577	3	US-08-658-136-1
c 7	47	1.5	13158	2	US-08-687-080-105
c 8	45	1.4	153	2	US-08-849-701-2
c 9	45	1.4	891	4	US-09-247-155-141
c 10	45	1.4	1701	4	US-09-078-294-9
c 11	45	1.4	3267	2	US-08-257-963B-12
c 12	45	1.4	3267	4	US-08-367-841A-12
c 13	45	1.4	3267	5	PCT-US95-07201-12
c 14	45	1.4	3663	4	US-09-499-884-11
c 15	45	1.4	5581	4	US-08-973-544-1
c 16	45	1.4	8409	4	US-09-167-681-37
c 17	45	1.4	22481	4	US-08-367-841A-43
c 18	45	1.4	22481	5	PCT-US95-07201-43
c 19	45	1.4	84495	4	US-09-797-906-3
c 20	45	1.4	246240	2	US-08-724-394A-20
c 21	45	1.4	246240	2	US-08-724-394A-21
c 22	45	1.4	246240	2	US-08-724-394A-22
c 23	44	1.4	176373	3	US-09-128-155-17
c 24	43	1.3	2310	1	US-08-471-570-9
c 25	43	1.3	2676	1	US-08-471-570-7
c 26	43	1.3	14796	4	US-08-975-080-35
c 27	43	1.3	14796	4	US-09-630-706-10

28 43 1.3 14796 4 US-09-496-694B-3 Sequence 3, Appli  
29 43 1.3 28720 4 US-09-341-587-7 Sequence 7, Appli  
30 43 1.3 72604 4 US-09-268-992-7 Sequence 7, Appli  
31 42 1.3 1613 2 US-08-812-204-1 Sequence 1, Appli  
32 42 1.3 1656 1 US-08-324-465-2 Sequence 2, Appli  
33 42 1.3 1656 2 US-08-465-981-2 Sequence 2, Appli  
34 42 1.3 1656 5 PCT-US93-11915-2 Sequence 2, Appli  
35 42 1.3 1725 1 US-08-324-465-5 Sequence 5, Appli  
36 42 1.3 1725 2 US-08-465-981-5 Sequence 5, Appli  
37 42 1.3 1725 5 PCT-US93-11915-5 Sequence 5, Appli  
38 42 1.3 2086 2 US-08-655-640-5 Sequence 5, Appli  
39 42 1.3 5590 4 US-09-050-159-129 Sequence 129, App  
40 41 1.3 198 2 US-08-967-101-107 Sequence 107, App  
41 1.3 198 2 US-08-593-541-107 Sequence 107, App  
42 41 1.3 198 3 US-09-124-698-107 Sequence 107, App  
43 41 1.3 198 4 US-09-127-480-107 Sequence 107, App  
44 41 1.3 198 4 US-08-496-841C-107 Sequence 107, App  
45 41 1.3 386 2 US-08-967-101-103 Sequence 103, App

#### ALIGNMENTS

RESULT 1  
US-08-724-394A-20/c  
; Sequence 20, Application US/08724394A  
; Patent No. 5872237  
; GENERAL INFORMATION:  
; APPLICANT: Feder, John N.  
; APPLICANT: Kronmal, Gregory S.  
; APPLICANT: Lauer, Peter M.  
; APPLICANT: Ruddy, David A.  
; APPLICANT: Thomas, Winston  
; APPLICANT: Tsuchihashi, Zenta  
; APPLICANT: Wolff, Roger K.  
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el  
; TITLE OF INVENTION: Sequences and Antibodies Thereto  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94111-3834

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/724,394A  
FILING DATE: 01-OCT-1996  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitts, Renee A.  
REGISTRATION NUMBER: 35,136  
REFERENCE/DOCKET NUMBER: 017957-000100  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-576-0200  
TELEFAX: 415-576-0300  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 246240 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: cdna  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: 1..246240  
OTHER INFORMATION: /note= "HLA-H.CONTIG"  
US-08-724-394A-20

Query Match 1.7%; Score 55; DB 2; Length 246240;  
Best Local Similarity 100.0%; Pred. No. 5.1e-15;  
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1190 cccacctggcctcccaaaagtctgggattacaggcatgagccactgcgcccagc 1244  
|||||  
Db 14447 CCCACCTTGGCCTCCCAAAAGTCTGGGATTACAGGCATGAGCCACTGCGGCCAGC 14393

## RESULT 2

US-08-724-394A-21/c  
; Sequence 21, Application US/08724394A  
; Patent No. 5872237  
; GENERAL INFORMATION:  
; APPLICANT: Feder, John N.  
; APPLICANT: Kronmal, Gregory S.  
; APPLICANT: Lauer, Peter M.  
; APPLICANT: Ruddy, David A.  
; APPLICANT: Thomas, Winston  
; APPLICANT: Tsuchihashi, Zenta  
; APPLICANT: Wolff, Roger K.  
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el  
; TITLE OF INVENTION: Sequences and Antibodies Thereto  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; FILING DATE: 01-OCT-1996  
; APPLICATION NUMBER: US/08/724,394A  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fitts, Renee A.  
; REGISTRATION NUMBER: 35,136  
; REFERENCE/DOCKET NUMBER: 017957-000100  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-576-0200  
; TELEFAX: 415-576-0300  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 246240 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: cdna  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..246240  
; OTHER INFORMATION: /note= "HLA-H.CONTIG"  
US-08-724-394A-21

Query Match 1.7%; Score 55; DB 2; Length 246240;  
Best Local Similarity 100.0%; Pred. No. 5.1e-15;  
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1190 cccacctggcctcccaaaagtctgggattacaggcatgagccactgcgcccagc 1244  
|||||  
Db 14447 CCCACCTTGGCCTCCCAAAAGTCTGGGATTACAGGCATGAGCCACTGCGGCCAGC 14393

## RESULT 3

US-08-724-394A-22/c  
; Sequence 22, Application US/08724394A  
; Patent No. 5872237  
; GENERAL INFORMATION:  
; APPLICANT: Feder, John N.  
; APPLICANT: Kronmal, Gregory S.  
; APPLICANT: Lauer, Peter M.  
; APPLICANT: Ruddy, David A.  
; APPLICANT: Thomas, Winston  
; APPLICANT: Tsuchihashi, Zenta  
; APPLICANT: Wolff, Roger K.  
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el  
; TITLE OF INVENTION: Sequences and Antibodies Thereto  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; FILING DATE: 01-OCT-1996  
; APPLICATION NUMBER: US/08/724,394A  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fitts, Renee A.  
; REGISTRATION NUMBER: 35,136  
; REFERENCE/DOCKET NUMBER: 017957-000100  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-576-0200  
; TELEFAX: 415-576-0300  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 246240 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: cdna  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..246240  
; OTHER INFORMATION: /note= "HLA-H.CONTIG"  
US-08-724-394A-22

Query Match 1.7%; Score 55; DB 2; Length 246240;  
Best Local Similarity 100.0%; Pred. No. 5.1e-15;  
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1190 cccacctggcctcccaaaagtctgggattacaggcatgagccactgcgcccagc 1244  
|||||  
Db 14447 CCCACCTTGGCCTCCCAAAAGTCTGGGATTACAGGCATGAGCCACTGCGGCCAGC 14393

## RESULT 4

US-08-323-443B-1  
; Sequence 1, Application US/08323443B  
; Patent No. 5654170  
; GENERAL INFORMATION:  
; APPLICANT: KLINGER, KATHERINE W.  
; APPLICANT: LANDES, GREGORY M.  
; APPLICANT: BURN, TIMOTHY C.  
; APPLICANT: CONNORS, TIMOTHY D.  
; APPLICANT: DACKOWSKI, WILLIAM R.  
; APPLICANT: GERMINO, GREGORY  
; APPLICANT: QIAN, FENG  
; TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENE

; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Darby & Darby PC  
; STREET: 805 Third Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10022  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; APPLICATION NUMBER: US/08/323,443B  
; FILING DATE: 12-OCT-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ludwig, S. Peter  
; REGISTRATION NUMBER: 25,351  
; REFERENCE/DOCKET NUMBER: 0372/0A462  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 527-7700  
; TELEFAX: (212) 753-6237  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 31571 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; IMMEDIATE SOURCE:  
; CLONE: PKD1 GENOMIC  
US-08-323-443B-1

Query Match 1.5%; Score 48; DB 1; Length 31571;  
Best Local Similarity 100.0%; Pred. No. 6.5e-12;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1187 ctgccacattggctcccaagtgctggattacaggcatgagccac 1234  
|||||  
DB 1117 CTGCCACCTTGGCTCCCAAGTCTGGATTACAGGCATGAGCCAC 1164

RESULT 5  
US-08-658-136-2  
; Sequence 2, Application US/08658136  
; Patent No. 6071717  
; GENERAL INFORMATION:  
; APPLICANT: KLINGER, KATHERINE W  
; APPLICANT: LANDES, GREGORY M  
; APPLICANT: BURN, TIMOTHY C  
; APPLICANT: CONNORS, TIMOTHY D  
; APPLICANT: DACKOWSKI, WILLIAM  
; APPLICANT: GERMINO, GREGORY  
; APPLICANT: QIAN, FENG  
; TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENE  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENZYME CORPORATION  
; STREET: ONE MOUNTAIN ROAD  
; CITY: FRAMINGHAM  
; STATE: MASSACHUSETTS  
; COUNTRY: USA  
; ZIP: 01701  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/658,136  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: LASSEN, ELIZABETH  
; REGISTRATION NUMBER: 31,845  
; REFERENCE/DOCKET NUMBER: GEN4-17.8  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 508-872-8400  
; TELEFAX: 508-872-5415  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 53526 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-658-136-2

Query Match 1.5%; Score 48; DB 3; Length 53526;  
Best Local Similarity 100.0%; Pred. No. 6.3e-12;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1187 ctgccacattggctcccaagtgctggattacaggcatgagccac 1234  
|||||  
DB 1712 CTGCCACCTTGGCTCCCAAGTCTGGATTACAGGCATGAGCCAC 1759

RESULT 6  
US-08-658-136-1  
; Sequence 1, Application US/08658136  
; Patent No. 6071717  
; GENERAL INFORMATION:  
; APPLICANT: KLINGER, KATHERINE W  
; APPLICANT: LANDES, GREGORY M  
; APPLICANT: BURN, TIMOTHY C  
; APPLICANT: CONNORS, TIMOTHY D  
; APPLICANT: DACKOWSKI, WILLIAM  
; APPLICANT: GERMINO, GREGORY  
; APPLICANT: QIAN, FENG  
; TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENE  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENZYME CORPORATION  
; STREET: ONE MOUNTAIN ROAD  
; CITY: FRAMINGHAM  
; STATE: MASSACHUSETTS  
; COUNTRY: USA  
; ZIP: 01701  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/658,136  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: LASSEN, ELIZABETH  
; REGISTRATION NUMBER: 31,845  
; REFERENCE/DOCKET NUMBER: GEN4-17.8  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 508-872-8400  
; TELEFAX: 508-872-5415  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 53577 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single

```
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-658-136-1

Query Match      1.5%; Score 48; DB 3; Length 53577;
Best Local Similarity 100.0%; Pred. No. 6.3e-12;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1187 ctgccacctggcctcccaagtgctgggattacagcatgagccac 1234
|||||
Db 1712 CTGCCACCTTGGCTCCCAAGTGCTGGGATTACAGGCATGAGCCAC 1759

RESULT 7
US-08-687-080-105
; Sequence 105, Application US/08687080
; Patent No. 5965427
; GENERAL INFORMATION:
; APPLICANT: Gregory Dolganov
; TITLE OF INVENTION: Human RAD50 Gene and Methods of Use Thereof
; NUMBER OF SEQUENCES: 175
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/687,080
; FILING DATE: 17-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/592,136
; FILING DATE: 26-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0111.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 105:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13158 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: 5' END OF INTRON 21 OF RAD50 GENOMIC
; INDIVIDUAL ISOLATE: SEQUENCE
US-08-687-080-105

Query Match      1.5%; Score 47; DB 2; Length 13158;
Best Local Similarity 100.0%; Pred. No. 1.9e-11;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1198 ggcctcccaagtgctgggattacagcatgagccactgcgccagc 1244
|||||
Db 336 GGCTCCCAAGTGCTGGGATTACAGGCATGAGCCACTGCCCCAGC 382

RESULT 8
```

```
US-08-849-701-2/c
; Sequence 2, Application US/08849701
; Patent No. 5922544
; GENERAL INFORMATION:
; APPLICANT: Miyai, Kiyoshi
; APPLICANT: Naitoh, Tsutomu
; APPLICANT: Yonekawa, Toshihiro
; TITLE OF INVENTION: Method of Cell Detection
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/849,701
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/JP95/02734
; FILING DATE: 27-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: EIKENI.001APC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 714-760-0404
; TELEFAX: 714-760-9502
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 153 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: Alu sequence BLUR1
US-08-849-701-2

Query Match      1.4%; Score 45; DB 2; Length 153;
Best Local Similarity 100.0%; Pred. No. 1.7e-10;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1190 ccacacttgccctcccaagtgctgggattacagcatgagccac 1234
|||||
Db 85 CCACCTTGGCTCCCAAGTGCTGGGATTACAGGCATGAGCCAC 41

RESULT 9
US-09-247-155-141/c
; Sequence 141, Application US/09247155A
; Patent No. 6312922
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean-Baptiste
; APPLICANT: Duclert, Aymeric
; APPLICANT: Bouqueleret, Lydie
; TITLE OF INVENTION: Complementary DNAs
; FILE REFERENCE: GENSET 021A
; CURRENT APPLICATION NUMBER: US/09/247,155A
; CURRENT FILING DATE: 1999-02-09
; EARLIER APPLICATION NUMBER: 60/074,121
; EARLIER FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: 60/081,563
; EARLIER FILING DATE: 1998-04-13
; EARLIER APPLICATION NUMBER: 60/096,116
```

; EARLIER FILING DATE: 1998-08-10  
 ; EARLIER APPLICATION NUMBER: 60/099,273  
 ; EARLIER FILING DATE: 1998-10-04  
 ; NUMBER OF SEQ ID NOS: 182  
 ; SOFTWARE: Patent.pm  
 ; SEQ ID NO 141  
 ; LENGTH: 891  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 4..447  
 ; FEATURE:  
 ; NAME/KEY: sig\_peptide  
 ; LOCATION: 4..147  
 ; OTHER INFORMATION: Von Heijne matrix  
 ; OTHER INFORMATION: score 5.6999980926514  
 ; OTHER INFORMATION: seq LLLFFGKLLVVG/VG  
 ; FEATURE:  
 ; NAME/KEY: polyA\_signal  
 ; LOCATION: 858..863  
 ; FEATURE:  
 ; NAME/KEY: polyA\_site  
 ; LOCATION: 880..891  
 ; US-09-247-155-141

Query Match 1.4%; Score 45; DB 4; Length 891;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-10;  
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1033 aacctgtcctccgggtcaagcattctctgcctcagcctcc 1077  
 |||||  
 Db 772 AACCTCTCCTCCGGGTTCAGCGATTCTCCTCAGCCCTCC 728

RESULT 10  
 US-09-078-294-9  
 ; Sequence 9, Application US/09078294  
 ; Patent No. 6265211  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Choo, Kong-Hong Andy  
 ; APPLICANT: Du Sart, Desiree  
 ; APPLICANT: Cancilla, Michael R.  
 ; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE  
 ; FILE REFERENCE: Davies Col  
 ; CURRENT APPLICATION NUMBER: US/09/078,294  
 ; CURRENT FILING DATE: 1998-05-13  
 ; NUMBER OF SEQ ID NOS: 29  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 9  
 ; LENGTH: 1701  
 ; TYPE: DNA  
 ; ORGANISM: BAC-F2 contig 5  
 ; US-09-078-294-9

Query Match 1.4%; Score 45; DB 4; Length 1701;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-10;  
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1033 aacctgtcctccgggtcaagcattctctgcctcagcctcc 1077  
 |||||  
 Db 1182 aacctgtcctccgggtcaagcattctctgcctcagcctcc 1226

RESULT 11  
 US-08-257-963B-12  
 ; Sequence 12, Application US/08257963B  
 ; Patent No. 5840686  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Chader, Gerald J.; Becerra, S.  
 ; APPLICANT: Patricia; Schwartz, Joan P.;

; APPLICANT: Taniwaki, Takayuki  
 ; TITLE OF INVENTION: PIGMENT EPITHELIUM  
 ; TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION OF ITS NOVEL  
 ; TITLE OF INVENTION: BIOLOGICAL ACTIVITY AND SEQUENCES ENCODING  
 ; TITLE OF INVENTION: AND EXPRESSING THE PROTEIN  
 ; NUMBER OF SEQUENCES: 42  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Morgan & Finnegan  
 ; STREET: 345 Park Avenue  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10154  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy Disk  
 ; COMPUTER: IBM PC Compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: WORDPERFECT 5.1  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/257,963B  
 ; FILING DATE:  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 07/952,796  
 ; FILING DATE: 24-SEPT-1992  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: DOROTHY R. AUTH  
 ; REGISTRATION NUMBER: 36434  
 ; REFERENCE/DOCKET NUMBER: 202664126US1  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 758-4800  
 ; TELEFAX: (212) 751-6849  
 ; INFORMATION FOR SEQ ID NO: 12:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 3267 Base Pairs  
 ; TYPE: Nucleic Acid  
 ; STRANDEDNESS: Double  
 ; TOPOLOGY: Unknown  
 ; MOLECULE TYPE: Genomic DNA  
 ; FEATURE:  
 ; NAME/KEY: JT109  
 ; LOCATION:  
 ; IDENTIFICATION METHOD:  
 ; OTHER INFORMATION: 3.3 kb PCR product  
 ; OTHER INFORMATION: using primers, SEQ ID No. 5840686 15 and 16  
 ; US-08-257-963B-12

Query Match 1.4%; Score 45; DB 2; Length 3267;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-10;  
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1190 cccacattggctcccaagtgctggattacagcattagccac 1234  
 |||||  
 Db 613 CCCACCTTGGCTCCCAAGTGTGGATTACAGCATGAGCCAC 657

RESULT 12  
 US-08-367-841A-12  
 ; Sequence 12, Application US/08367841A  
 ; Patent No. 6319687  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Chader, Gerald J.; Rodriguez,  
 ; APPLICANT: Ignacio R.; Mazuruk, Krzysztof;  
 ; APPLICANT: Tombran-Tink, Joyce  
 ; TITLE OF INVENTION: PIGMENT EPITHELIUM  
 ; TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION GENOMIC  
 ; TITLE OF INVENTION: ORGANIZATION AND SEQUENCE OF THE PEDF GENE  
 ; NUMBER OF SEQUENCES: 43  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Morgan & Finnegan  
 ; STREET: 345 Park Avenue  
 ; CITY: New York

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07201
; FILING DATE: 06-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/367,841
; FILING DATE: 30-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/257,963
; FILING DATE: 07-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/952,796
; FILING DATE: 24-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: DOROTHY R. AUTH
; REGISTRATION NUMBER: 36434
; REFERENCE/DOCKET NUMBER: 20264126PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3267 Base Pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Double
; TOPOLOGY: Unknown
; MOLECULE TYPE: Genomic DNA
; FEATURE:
; NAME/KEY: JT109
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: 3.3 kb PCR product
; OTHER INFORMATION: using primers, SEQ ID No: 15 and 16
PCT-US95-07201-12

Query Match 1.4%; Score 45; DB 5; Length 3267;
Best Local Similarity 100.0%; Pred. No. 1.5e-10; Mismatches 0; Indels 0; Gaps 0;
Matches 45; Conservative 0;

QY 1190 ccaacttgctcccccagaagtctggattacagcgtagccac 1234
      |||||||||||||||||||||||||||||||||||||||
Db 613 CCCACCTGGCTCCCAAGTCTGGGATTACAGGCATGAGCCAC 657

RESULT 14
US-09-499-884-11/c
; Sequence 11, Application US/09499884
; Patent No. 6265172
; GENERAL INFORMATION:
; APPLICANT: St. Clair, Daret
; APPLICANT: Urano, Muneyasu
; APPLICANT: Kasarskis, Edward
; TITLE OF INVENTION: DIAGNOSTIC TEST AND THERAPY FOR MANGANESE SUPEROXIDE DISMUTASE
; TITLE OF INVENTION: ASSOCIATED DISEASES
; FILE REFERENCE: 50229-180
; CURRENT APPLICATION NUMBER: US/09/499,884
; CURRENT FILING DATE: 2000-02-08
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 3663
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-499-884-11

Query Match 1.4%; Score 45; DB 4; Length 3663;
Best Local Similarity 100.0%; Pred. No. 1.5e-10;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1033 aacctgctcccggttcaagcgtattcctgctcagctccc 1077
      |||||||||||||||||||||||||||||||||||||||

```



Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1033 aaactctgctcccggttcaagcgattctctgctcagctcc 1077  
|||||  
Db 4362 AACCTGCTGCTCCGGGTTCAAGCGATTCTCTGCTCAGCCTCC 4318

RESULT 16

US-09-167-681-37  
; Sequence 37, Application US/09167681A  
; Patent No. 6265561  
; GENERAL INFORMATION:  
; APPLICANT: Weinshilboum, M.D., Richard M.  
; APPLICANT: Raftogianis, Rebecca B.  
; APPLICANT: Wood, Thomas C.  
; APPLICANT: Otterness, Diane M.  
; TITLE OF INVENTION: SULFOTRANSFERASE SEQUENCE VARIANTS  
; FILE REFERENCE: 07039/118001  
; CURRENT APPLICATION NUMBER: US/09/167,681A  
; CURRENT FILING DATE: 1998-10-07  
; NUMBER OF SEQ ID NOS: 52  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 37  
; LENGTH: 8397  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (3730)...(3879)  
; NAME/KEY: CDS  
; LOCATION: (3987)...(4112)  
; NAME/KEY: CDS  
; LOCATION: (4198)...(4293)  
; NAME/KEY: CDS  
; LOCATION: (6088)...(6213)  
; NAME/KEY: CDS  
; LOCATION: (6309)...(6404)  
; NAME/KEY: CDS  
; LOCATION: (7214)...(7393)  
; NAME/KEY: CDS  
; LOCATION: (7516)...(7629)  
US-09-167-681-37

Query Match 1.4%; Score 45; DB 4; Length 8409;  
Best Local Similarity 100.0%; Pred. No. 1.4e-10;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1033 aaactctgctcccggttcaagcgattctctgctcagctcc 1077  
|||||  
Db 701 aaactctgctcccggttcaagcgattctctgctcagctcc 745

RESULT 17

US-08-367-841A-43  
; Sequence 43, Application US/08367841A  
; Patent No. 6319687  
; GENERAL INFORMATION:  
; APPLICANT: Chager, Gerald J.; Rodriguez,  
; APPLICANT: Ignacio R.; Mazuruk, Krzysztof;  
; APPLICANT: Tombran-Tink, Joyce  
; TITLE OF INVENTION: PIGMENT EPITHELIUM  
; TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION GENOMIC  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Morgan & Finnegan  
; STREET: 345 Park Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10154  
; COMPUTER READABLE FORM:

Db 1970 AACCTGCTGCTCCGGGTTCAAGCGATTCTCTGCTCAGCCTCC 1926

RESULT 15

US-08-973-544-1/C  
; Sequence 1, Application US/08973544  
; Patent No. 6338950  
; GENERAL INFORMATION:  
; APPLICANT: WEISS, Elisabeth  
; TITLE OF INVENTION: NEW IMMUNOREGULATORY PROTEIN LST-1  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIKAIKO, MARCELSTEIN, MURRAY & ORAM LLP  
; STREET: 655 Fifteenth St., NW, Suite 300, G St. Lobby  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-5701  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/973,544  
; FILING DATE: 18-DEC-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT EP 96/02663  
; FILING DATE: 20-JUN-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: EP 95109511.6  
; FILING DATE: 20-JUN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: EP 95112201.9  
; FILING DATE: 03-AUG-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kitts, Monica Chin  
; REGISTRATION NUMBER: 36,105  
; REFERENCE/DOCKET NUMBER: P8341-7073  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 638-5000  
; TELEFAX: (202) 638-4810  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5581 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: 48..162  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: 544..652  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: 1044..1162  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: 1475..1567  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: 1775..1797  
; FEATURE:  
; NAME/KEY: exon  
; LOCATION: 2325..2709  
US-08-973-544-1

Query Match 1.4%; Score 45; DB 4; Length 5581;  
Best Local Similarity 100.0%; Pred. No. 1.5e-10;

```
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/367,841A
; FILING DATE: 30-DEC-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/257,963
; FILING DATE: 07-JUN-1994
; APPLICATION DATA:
; APPLICATION NUMBER: 07/952,796
; FILING DATE: 24-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: DOROTHY R. AUTH
; REGISTRATION NUMBER: 36434
; REFERENCE/DOCKET NUMBER: 20264126US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22481 Base Pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Double
; TOPOLOGY: Unknown
; MOLECULE TYPE: Genomic DNA
; FEATURE:
; NAME/KEY: Pl-147
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: full length genomic
; OTHER INFORMATION: sequence for PEDF plus flanking sequences.
;
PCT-US95-07201-43
```

```
Query Match 1.4%; Score 45; DB 4; Length 22481;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1190 cccacctggcctcccaaaagtgtggattacagcagcatgagccac 1234
|||||
Db 17174 CCCACCTGGCCTCCCAAAAGTGTGGATTACAGGCATGAGCCAC 17218
```

```
RESULT 18
PCT-US95-07201-43
; Sequence 43, Application PC/TUS9507201
; GENERAL INFORMATION:
; APPLICANT: Chader, Gerald J.; Becerra, Sofia
; APPLICANT: Patricia; Schwartz, Joan P.;
; APPLICANT: Taniwaki, Takayuki
; TITLE OF INVENTION: PIGMENT EPITHELIUM
; TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION GENOMIC
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan, L.L.P.
; STREET: 345 Park Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07201
; FILING DATE: 06-JUN-1995
; CLASSIFICATION:
```

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/367,841
; FILING DATE: 30-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/257,963
; FILING DATE: 07-JUN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/952,796
; FILING DATE: 24-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: DOROTHY R. AUTH
; REGISTRATION NUMBER: 36434
; REFERENCE/DOCKET NUMBER: 20264126PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22481 Base Pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Double
; TOPOLOGY: Unknown
; MOLECULE TYPE: Genomic DNA
; FEATURE:
; NAME/KEY: Pl-147
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: full length genomic
; OTHER INFORMATION: sequence for PEDF plus flanking sequences.
;
PCT-US95-07201-43
```

```
Query Match 1.4%; Score 45; DB 5; Length 22481;
Best Local Similarity 100.0%; Pred. No. 1.4e-10;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1190 cccacctggcctcccaaaagtgtggattacagcagcatgagccac 1234
|||||
Db 17174 CCCACCTGGCCTCCCAAAAGTGTGGATTACAGGCATGAGCCAC 17218
```

```
RESULT 19
US-09-797-906-3
; Sequence 3, Application US/09797906
; Patent No. 6329188
; GENERAL INFORMATION:
; APPLICANT: Zianghe YAN, Karen A. KETCHUM, Valentina DIFRANCESCO, Ellen M. BEASLEY
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CLO01151CIP
; CURRENT APPLICATION NUMBER: US/09/797,906
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 84495
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(84495)
; OTHER INFORMATION: n = A,T,C or G
;
US-09-797-906-3
```

```
Query Match 1.4%; Score 45; DB 4; Length 84495;
Best Local Similarity 100.0%; Pred. No. 1.3e-10;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1192 caccttggcctcccaaaagtgtggattacagcagcatgagccactg 1236
|||||
Db 26148 caccttggcctcccaaaagtgtggattacagcagcatgagccactg 26192
```

## RESULT 20

US-08-724-394A-20  
; Sequence 20, Application US/08724394A  
; Patent No. 5872237

## GENERAL INFORMATION:

; APPLICANT: Feder, John N.  
; APPLICANT: Kronmal, Gregory S.  
; APPLICANT: Lauer, Peter M.  
; APPLICANT: Ruddy, David A.  
; APPLICANT: Thomas, Winston  
; APPLICANT: Tsuchihashi, Zenta  
; APPLICANT: Wolff, Roger K.

; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el

; TITLE OF INVENTION: Sequences and Antibodies thereto

; NUMBER OF SEQUENCES: 31

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP

; STREET: Two Embarcadero Center, 8th Floor

; CITY: San Francisco

; STATE: CA

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/724,394A

; FILING DATE: 01-OCT-1996

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: Fitts, Renee A.

; REGISTRATION NUMBER: 35,136

; REFERENCE/DOCKET NUMBER: 017957-000100

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-576-0200

; TELEFAX: 415-576-0300

; INFORMATION FOR SEQ ID NO: 20:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 246240 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: not relevant

; TOPOLOGY: not relevant

; MOLECULE TYPE: cdna

; FEATURE:

; NAME/KEY: misc.feature

; LOCATION: 1..246240

; OTHER INFORMATION: /note= "HLA-H.CONTIG"

US-08-724-394A-20

## Query Match

1.4%; Score 45; DB 2; Length 246240;

Best Local Similarity 100.0%; Pred. No. 1.2e-10;

Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1033 aacctctgctccgggttaacgagattctctgctcagctcc 1077

|||||

Db 242021 AACCTCTGCTCCGGGTTCAGCGATTCTCCTGCTCAGCTCC 242065

## RESULT 21

US-08-724-394A-21

; Sequence 21, Application US/08724394A

; Patent No. 5872237

## GENERAL INFORMATION:

; APPLICANT: Feder, John N.

; APPLICANT: Kronmal, Gregory S.

; APPLICANT: Lauer, Peter M.

; APPLICANT: Ruddy, David A.

; APPLICANT: Thomas, Winston

; APPLICANT: Tsuchihashi, Zenta

; APPLICANT: Wolff, Roger K.

; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el

; TITLE OF INVENTION: Sequences and Antibodies Thereto

; NUMBER OF SEQUENCES: 31

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP

; STREET: Two Embarcadero Center, 8th Floor

; CITY: San Francisco

; STATE: CA

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/724,394A

; FILING DATE: 01-OCT-1996

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: Fitts, Renee A.

; REGISTRATION NUMBER: 35,136

; REFERENCE/DOCKET NUMBER: 017957-000100

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-576-0200

; TELEFAX: 415-576-0300

; INFORMATION FOR SEQ ID NO: 21:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 246240 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: not relevant

; TOPOLOGY: not relevant

; MOLECULE TYPE: cdna

; FEATURE:

; NAME/KEY: misc.feature

; LOCATION: 1..246240

; OTHER INFORMATION: /note= "HLA-H.CONTIG"

US-08-724-394A-21

## Query Match

1.4%; Score 45; DB 2; Length 246240;

Best Local Similarity 100.0%; Pred. No. 1.2e-10;

Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1033 aacctctgctccgggttaacgagattctctgctcagctcc 1077

|||||

Db 242021 AACCTCTGCTCCGGGTTCAGCGATTCTCCTGCTCAGCTCC 242065

## RESULT 22

US-08-724-394A-22

; Sequence 22, Application US/08724394A

; Patent No. 5872237

## GENERAL INFORMATION:

; APPLICANT: Feder, John N.

; APPLICANT: Kronmal, Gregory S.

; APPLICANT: Lauer, Peter M.

; APPLICANT: Ruddy, David A.

; APPLICANT: Thomas, Winston

; APPLICANT: Tsuchihashi, Zenta

; APPLICANT: Wolff, Roger K.

; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el

; TITLE OF INVENTION: Sequences and Antibodies thereto

; NUMBER OF SEQUENCES: 31

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP

; STREET: Two Embarcadero Center, 8th Floor

; CITY: San Francisco

; STATE: CA

; COUNTRY: USA

; ZIP: 94111-3834

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/724,394A  
FILING DATE: 01-OCT-1996  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Pitts, Renee A.  
REGISTRATION NUMBER: 35,136  
REFERENCE/DOCKET NUMBER: 017957-000100  
TELEPHONE: 415-576-0200  
TELEFAX: 415-576-0300  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 246240 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1..246240  
OTHER INFORMATION: /note= "HLA-H. CONTIG"  
US-08-724-394A-22

Query Match 1.4%; Score 45; DB 2; Length 246240;  
Best Local Similarity 100.0%; Pred. No. 1.2e-10;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1033 aacctgtctcccggttcaagcgattctcgtccacgctcc 1077  
|||||  
Db 242021 AACCTCTGCTCCGGGTCAAGCGATTCTCCTGCCTCACCTCC 242065

RESULT 23  
US-09-128-155-17/c  
Sequence 17, Application US/09128155  
Patent No. 6117654  
GENERAL INFORMATION:  
APPLICANT: Pan, Yang  
TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY  
FILE OF INVENTION: AND USES THEREOF  
FILE REFERENCE: 09404/052001  
CURRENT APPLICATION NUMBER: US/09/128,155  
CURRENT FILING DATE: 1998-08-03  
EARLIER APPLICATION NUMBER: US 60/091,650  
EARLIER FILING DATE: 1998-07-02  
EARLIER APPLICATION NUMBER: US 60/054,646  
EARLIER FILING DATE: 1997-08-04  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 17  
LENGTH: 176373  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(176373)  
OTHER INFORMATION: n = A,T,C or G  
US-09-128-155-17

Query Match 1.4%; Score 44; DB 3; Length 176373;  
Best Local Similarity 100.0%; Pred. No. 3.4e-10;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1194 ccttgccctcccaagtctggattacagcagcactgc 1237  
|||||

Db 45158 CCTTGGCCTCCCAAGTCTGGGATTACAGGCATGAGCCACTGC 45115  
RESULT 24  
US-08-471-570-9/c  
Sequence 9, Application US/08471570  
Patent No. 5750371  
GENERAL INFORMATION:  
APPLICANT: IGARASHI, Koichi  
APPLICANT: SENOO, Masaharu  
APPLICANT: WATANABE, Tatsuya  
TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &  
ADDRESS: CUSHMAN  
STREET: 130 Water Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: US  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/471,570  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/149,664  
FILING DATE:  
APPLICATION NUMBER: US 07/743369  
FILING DATE: 16-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: LINEK, Ernest V  
REGISTRATION NUMBER: 29822  
REFERENCE/DOCKET NUMBER: 40897  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)523-3400  
TELEFAX: (617)523-6440  
TELEX: 200291 STRE UR  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2310 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 25..1980  
US-08-471-570-9

Query Match 1.3%; Score 43; DB 1; Length 2310;  
Best Local Similarity 100.0%; Pred. No. 1.1e-09;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1198 ggcctcccaagtctggattacagcagcactgcgcc 1240  
|||||  
Db 2066 GGCCTCCCAAGTCTGGGATTACAGGCATGAGCCACTGC GCC 2024

RESULT 25  
US-08-471-570-7/c  
Sequence 7, Application US/08471570  
Patent No. 5750371  
GENERAL INFORMATION:  
APPLICANT: IGARASHI, Koichi  
APPLICANT: SENOO, Masaharu  
APPLICANT: WATANABE, Tatsuya

;; TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF  
;; NUMBER OF SEQUENCES: 18  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &  
;; ADDRESSEE: CUSHMAN  
;; STREET: 130 Water Street  
;; CITY: Boston  
;; STATE: Massachusetts  
;; COUNTRY: US  
;; ZIP: 02109  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/471,570  
;; FILING DATE: 06-JUN-1995  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/149,664  
;; FILING DATE:  
;; APPLICATION NUMBER: US 07/743369  
;; FILING DATE: 16-AUG-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: LINEK, Ernest V  
;; REGISTRATION NUMBER: 29822  
;; REFERENCE/DOCKET NUMBER: 40897  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617)523-3400  
;; TELEFAX: (617)523-6440  
;; TELEX: 200291 STRE UR  
;; INFORMATION FOR SEQ ID NO: 7:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 2676 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: double  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: cDNA  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 25..2331  
US-08-471-570-7  
  
Query Match 1.3%; Score 43; DB 1; Length 2676;  
Best Local Similarity 100.0%; Pred. No. 1.1e-09;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1198 ggcctcccaagctgctggattacagcgatgagccactgcgcc 1240  
Db 2417 ggcctcccaagctgctggattacagcgatgagccactgcgcc 2375  
|||||  
  
RESULT 26  
US-08-975-080-35  
; Sequence 35, Application US/08975080  
; Patent No. 6245523  
; GENERAL INFORMATION:  
; APPLICANT: Altieri, Dario C.  
; TITLE OF INVENTION: SURVIVIN, A PROTEIN THAT INHIBITS  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP  
; STREET: 1800 M Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20036-5869  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/975,080  
;; FILING DATE: 20-NOV-1997  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 60/031,435  
;; FILING DATE: 20-NOV-1996  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Adler, Reid G.  
;; REGISTRATION NUMBER: 30,988  
;; REFERENCE/DOCKET NUMBER: 044574-5022-01-WO  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 202-467-7000  
;; TELEFAX: 202-467-7176  
;; INFORMATION FOR SEQ ID NO: 35:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 14796 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: DNA (genomic)  
US-08-975-080-35  
  
Query Match 1.3%; Score 43; DB 4; Length 14796;  
Best Local Similarity 100.0%; Pred. No. 1e-09;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1035 cctctgctcccggttcaagcgattctctgctcagctcc 1077  
Db 2320 CCTCTGCTCCCGGTTCAAGCGATTCTCTGCTCAGCCTCC 2362  
|||||

RESULT 27  
US-09-630-706-10  
; Sequence 10, Application US/09630706  
; Patent No. 6277640  
; GENERAL INFORMATION:  
; APPLICANT: C. Frank Bennett  
; APPLICANT: Lex M. Cowser  
; TITLE OF INVENTION: ANTISENSE MODULATION OF HER-3 EXPRESSION  
; FILE REFERENCE: RTS-0053  
; CURRENT APPLICATION NUMBER: US/09/630,706  
; CURRENT FILING DATE: 2000-08-01  
; NUMBER OF SEQ ID NOS: 94  
; SEQ ID NO 10  
; LENGTH: 14796  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (2811)...(2921)  
; NAME/KEY: CDS  
; LOCATION: (3174)...(3283)  
; NAME/KEY: CDS  
; LOCATION: (5158)...(5275)  
; NAME/KEY: CDS  
; LOCATION: (11955)...(12044)  
US-09-630-706-10

Query Match 1.3%; Score 43; DB 4; Length 14796;  
Best Local Similarity 100.0%; Pred. No. 1e-09;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1035 cctctgctcccggttcaagcgattctctgctcagctcc 1077  
Db 2320 cctctgctcccggttcaagcgattctctgctcagctcc 2362  
|||||

RESULT 28  
US-09-496-694B-3

; Sequence 3, Application US/09496694B  
; Patent No. 6335194  
; GENERAL INFORMATION:  
; APPLICANT: C. Frank Bennett  
; APPLICANT: Elizabeth J. Ackermann  
; APPLICANT: Eric E. Swayze  
; APPLICANT: Lex M. Cowser  
; TITLE OF INVENTION: ANTISENSE MODULATION OF SURVIVIN EXPRESSION  
; FILE REFERENCE: ISPH-0439  
; CURRENT APPLICATION NUMBER: US/09/496,694B  
; PRIOR FILING DATE: 2000-02-02  
; PRIOR APPLICATION NUMBER: 09/286,407  
; PRIOR FILING DATE: 1999-04-05  
; PRIOR APPLICATION NUMBER: 09/163,162  
; PRIOR FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 249  
; SEQ ID NO 3  
; LENGTH: 14796  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (2811)...(2921)  
; NAME/KEY: CDS  
; LOCATION: (3174)...(3283)  
; NAME/KEY: CDS  
; LOCATION: (5158)...(5275)  
; NAME/KEY: CDS  
; LOCATION: (11955)...(12044)  
US-09-496-694B-3

Query Match 1.3%; Score 43; DB 4; Length 14796;  
Best Local Similarity 100.0%; Pred. No. 1e-09;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1035 cctctgctccgggttaacgagcttcctcgtcagctccc 1077  
|||||  
Db 2320 cctctgctccgggttaacgagcttcctcgtcagctccc 2362

## RESULT 29

US-09-341-587-7  
; Sequence 7, Application US/09341587  
; Patent No. 6346606  
; GENERAL INFORMATION:  
; APPLICANT: Mollenhauer, Jan  
; TITLE OF INVENTION: Protein Containing an SRCR Domain  
; FILE REFERENCE: 4121-108  
; CURRENT APPLICATION NUMBER: US/09/341,587  
; CURRENT FILING DATE: 1999-08-31  
; EARLIER APPLICATION NUMBER: PCT/DE98/00096  
; EARLIER FILING DATE: 1998-01-09  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 28720  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-341-587-7

Query Match 1.3%; Score 43; DB 4; Length 28720;  
Best Local Similarity 100.0%; Pred. No. 1e-09;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1033 aacctgctccgggttaacgagcttcctcgtcagctccc 1075  
|||||  
Db 19480 aacctgctccgggttaacgagcttcctcgtcagctccc 19522

## RESULT 30

US-09-268-992-7/c

; Sequence 7, Application US/09268992  
; Patent No. 6342351  
; GENERAL INFORMATION:  
; APPLICANT: Chen, H.  
; APPLICANT: Freimer, N.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING  
; FILE REFERENCE: 7853-138  
; CURRENT APPLICATION NUMBER: US/09/268,992  
; CURRENT FILING DATE: 1999-03-16  
; EARLIER APPLICATION NUMBER: 09/236,134  
; EARLIER FILING DATE: 1999-01-22  
; EARLIER APPLICATION NUMBER: 60/106,056  
; EARLIER FILING DATE: 1998-10-28  
; EARLIER APPLICATION NUMBER: 60/088,312  
; EARLIER FILING DATE: 1998-06-05  
; EARLIER APPLICATION NUMBER: 60/078,044  
; EARLIER FILING DATE: 1998-03-16  
; NUMBER OF SEQ ID NOS: 84  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 7  
; LENGTH: 72604  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: all n positions  
; OTHER INFORMATION: n=a, c, g, or t  
US-09-268-992-7

Query Match 1.3%; Score 43; DB 4; Length 72604;  
Best Local Similarity 100.0%; Pred. No. 9.6e-10;  
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1199 gctctcccaagtgctgggttacagcagcatgagccactgcgcc 1241  
|||||  
Db 63856 GCCTCCCAAAGCTGGGATTACAGCATGAGCCACTGGCC 63814

## RESULT 31

US-08-812-204-1  
; Sequence 1, Application US/08812204  
; Patent No. 5965790  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan L.  
; TITLE OF INVENTION: SR-BI REGULATORY SEQUENCES AND  
; FILE REFERENCE: THERAPEUTIC METHODS OF USE  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
; STREET: One Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109-2170  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/812,204  
; FILING DATE: 06-MAR-1997  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Arnold, Beth E.  
; REGISTRATION NUMBER: 35,430  
; REFERENCE/DOCKET NUMBER: MIA-014.01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-832-1000  
; TELEFAX: 617-832-7000  
; INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
; LENGTH: 1613 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-812-204-1

Query Match 1.3%; Score 42; DB 2; Length 1613;  
Best Local Similarity 100.0%; Pred. No. 3.2e-09;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1040 gctcccggttcaagcgattctctgctcagctcctctgag 1081  
|||||  
Db 447 GCCTCCGGGTTCAAGCGATTCTCTGCTCAGCTCCTCTGAG 488

RESULT 33  
US-08-465-981-2  
; Sequence 2, Application US/08465981  
; Patent No. 5874415  
; GENERAL INFORMATION:  
; APPLICANT: Kufe, Donald  
; APPLICANT: Abe, Miyako  
; TITLE OF INVENTION: ENHANCER SEQUENCE FOR MODULATING  
; TITLE OF INVENTION: EXPRESSION IN EPITHELIAL CELLS  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; COMPUTER: IBM PS/2 Model 50Z or 55SX  
; OPERATING SYSTEM: MS-DOS (Version 5.0)  
; SOFTWARE: WordPerfect (Version 5.1)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/465,981  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/324,465  
; FILING DATE: October 17, 1994  
; APPLICATION NUMBER: 07/999,742  
; FILING DATE: December 31, 1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fraser, Janis K.  
; REGISTRATION NUMBER: 34,819  
; REFERENCE/DOCKET NUMBER: 00530/065002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1656  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-08-465-981-2

Query Match 1.3%; Score 42; DB 2; Length 1656;  
Best Local Similarity 100.0%; Pred. No. 3.2e-09;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1040 gctcccggttcaagcgattctctgctcagctcctctgag 1081  
|||||  
Db 332 GCCTCCGGGTTCAAGCGATTCTCTGCTCAGCTCCTCTGAG 373

RESULT 34  
PCT-US93-11915-2  
; Sequence 2, Application PC/TUS9311915  
; GENERAL INFORMATION:  
; APPLICANT: Kufe, Donald  
; APPLICANT: Abe, Miyako  
; TITLE OF INVENTION: ENHANCER SEQUENCE FOR MODULATING  
; TITLE OF INVENTION: EXPRESSION IN EPITHELIAL CELLS  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.

SEQUENCE CHARACTERISTICS:  
; LENGTH: 1613 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-812-204-1

Query Match 1.3%; Score 42; DB 2; Length 1613;  
Best Local Similarity 100.0%; Pred. No. 3.2e-09;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1040 gctcccggttcaagcgattctctgctcagctcctctgag 1081  
|||||  
Db 447 GCCTCCGGGTTCAAGCGATTCTCTGCTCAGCTCCTCTGAG 488

RESULT 32  
US-08-324-465-2  
; Sequence 2, Application US/08324465  
; Patent No. 5565334  
; GENERAL INFORMATION:  
; APPLICANT: Kufe, Donald  
; APPLICANT: Abe, Miyako  
; TITLE OF INVENTION: GENE TRANSCRIPTION AND  
; TITLE OF INVENTION: IONIZING RADIATION: METHODS  
; TITLE OF INVENTION: AND COMPOSITIONS  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; COMPUTER: IBM PS/2 Model 50Z or 55SX  
; OPERATING SYSTEM: MS-DOS (Version 5.0)  
; SOFTWARE: WordPerfect (Version 5.1)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/324,465  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/999,742  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fraser, Janis K.  
; REGISTRATION NUMBER: 34,819  
; REFERENCE/DOCKET NUMBER: 00530/065001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1656  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-08-324-465-2

Query Match 1.3%; Score 42; DB 1; Length 1656;  
Best Local Similarity 100.0%; Pred. No. 3.2e-09;  
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1040 gctcccggttcaagcgattctctgctcagctcctctgag 1081  
|||||  
Db 332 GCCTCCGGGTTCAAGCGATTCTCTGCTCAGCTCCTCTGAG 373

```

;
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/11915
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/999,742
; FILING DATE: December 31, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00530/065W01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1656
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; PCT-US93-11915-2

```

```

Query Match 1.3%; Score 42; DB 5; Length 1656;
Best Local Similarity 100.0%; Pred. No. 3.2e-09;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1040 gctccgggttcaagcgattctctgcctcagcctcctgag 1081
|||||
Db 332 GCCTCCGGGTTCAAGCGATTCTCTGCTCAGCCTCCTGAG 373

```

```

RESULT 35
US-08-324-465-5
; Sequence 5, Application US/08324465
; Patent No. 5565334
; GENERAL INFORMATION:
; APPLICANT: Kufe, Donald
; APPLICANT: Abe, Miyako
; TITLE OF INVENTION: GENE TRANSCRIPTION AND
; TITLE OF INVENTION: IONIZING RADIATION: METHODS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/324,465
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/999,742
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00530/065001

```

```

;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1725
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-324-465-5

```

```

Query Match 1.3%; Score 42; DB 1; Length 1725;
Best Local Similarity 100.0%; Pred. No. 3.1e-09;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1040 gctccgggttcaagcgattctctgcctcagcctcctgag 1081
|||||
Db 332 GCCTCCGGGTTCAAGCGATTCTCTGCTCAGCCTCCTGAG 373

```

```

RESULT 36
US-08-465-981-5
; Sequence 5, Application US/08465981
; Patent No. 5874415
; GENERAL INFORMATION:
; APPLICANT: Kufe, Donald
; APPLICANT: Abe, Miyako
; TITLE OF INVENTION: ENHANCER SEQUENCE FOR MODULATING
; TITLE OF INVENTION: EXPRESSION IN EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,981
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/324,465
; FILING DATE: October 17, 1994
; APPLICATION NUMBER: 07/999,742
; FILING DATE: December 31, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00530/065002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1725
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-465-981-5

```

```

Query Match 1.3%; Score 42; DB 2; Length 1725;
Best Local Similarity 100.0%; Pred. No. 3.1e-09;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```



```

; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/655,640
; FILING DATE: 30-MAY-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/992,538
; FILING DATE: December 17, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: SCOTT, WATSON T.
; REGISTRATION NUMBER: 26,581
; REFERENCE/DOCKET NUMBER: WTS/5683/98019/SAP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2086 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-655-640-5

```

```

Query Match 1.3%; Score 42; DB 2; Length 2086;
Best Local Similarity 100.0%; Pred. No. 3.le-09;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1195 cttggcctcccaagtgtgggattacagcgatgagccactg 1236
|||||
Db 192 CTTGGCCTCCCAAGTGTGGGATTACAGCGATGAGCCACTG 233
|||||

```

```

RESULT 39
US-09-050-159-129
; Sequence 129, Application US/09050159A
; Patent No. 6197505
; GENERAL INFORMATION:
; APPLICANT: No. 6197505berg, Leif T
; APPLICANT: Andersson, Maria K
; TITLE OF INVENTION: METHODS FOR ASSESSING CARDIOVASCULAR STATUS AND
; TITLE OF INVENTION: COMPOSITIONS FOR USE THEREOF
; FILE REFERENCE: 1248/1D042
; CURRENT APPLICATION NUMBER: US/09/050,159A
; EARLIER FILING DATE: 1998-03-27
; EARLIER APPLICATION NUMBER: 60/042,930
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 129
; LENGTH: 5590
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Angiotensin I converting enzyme, 5' region
; US-09-050-159-129

```

```

Query Match 1.3%; Score 42; DB 4; Length 5590;
Best Local Similarity 100.0%; Pred. No. 3e-09;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1036 cttgctcccggttcaagcgattctcctgcctcagcctcc 1077
|||||

```

```

Qy 1040 gctcccggttcaagcgattctcctgcctcagcctcctgag 1081
|||||
Db 332 GCTCCCGGTTCAAGCGATTCTCTGCTCAGCTCCTGAG 373
|||||
RESULT 37
PCT-US93-11915-5
; Sequence 5, Application PC/TUS9311915
; GENERAL INFORMATION:
; APPLICANT: Kufe, Donald
; APPLICANT: Abe, Miyako
; TITLE OF INVENTION: ENHANCER SEQUENCE FOR MODULATING
; TITLE OF INVENTION: EXPRESSION IN EPITHELIAL CELLS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/11915
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/999,742
; FILING DATE: December 31, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00530/065W01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1725
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; PCT-US93-11915-5

```

```

Query Match 1.3%; Score 42; DB 5; Length 1725;
Best Local Similarity 100.0%; Pred. No. 3.le-09;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1040 gctcccggttcaagcgattctcctgcctcagcctcctgag 1081
|||||
Db 332 GCTCCCGGTTCAAGCGATTCTCTGCTCAGCTCCTGAG 373
|||||

```

```

RESULT 38
US-08-655-640-5
; Sequence 5, Application US/08655640
; Patent No. 5948613
; GENERAL INFORMATION:
; APPLICANT: Teng, Christina
; APPLICANT: Panella, Timothy J.
; TITLE OF INVENTION: HUMAN LACTOPERRIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARRY & CUSHMAN
; STREET: 1100 NEW YORK AVE. N.W., NINTH FLOOR
; CITY: WASHINGTON

```

Db 1837 ctctgctccgggttaagcgattctctgctcagcctcc 1878

RESULT 40  
US-08-967-101-107  
; Sequence 107, Application US/08967101  
; Patent No. 5840540  
; GENERAL INFORMATION:  
; APPLICANT: ST. GEORGE-HYSLOP, PETER H  
; APPLICANT: ROMMENS, JOHANNA M  
; APPLICANT: FRASER, PAUL E  
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED  
; TITLE OF INVENTION: TO ALZHEIMER'S DISEASE  
; NUMBER OF SEQUENCES: 183  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT  
; STREET: High Street Tower - 125 High Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02110  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; FILING DATE: 10-NOV-1997  
; CLASSIFICATION: 435  
; APPLICATION NUMBER: 08/592,541  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pitcher, Edmund R.  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 107:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 198 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; US-08-967-101-107

Query Match 1.3%; Score 41; DB 2; Length 198;  
Best Local Similarity 100.0%; Pred. No. 9.6e-09;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1194 ccttgccctccaaagtctggattacagcgatgagccac 1234  
|||||  
Db 12 CCTTGGCCTCCCAAGTCTGGGATTACAGGCATGAGCCAC 52

RESULT 41  
US-08-592-541-107  
; Sequence 107, Application US/08592541  
; Patent No. 5986054  
; GENERAL INFORMATION:  
; APPLICANT: ST. GEORGE-HYSLOP, PETER H  
; APPLICANT: ROMMENS, JOHANNA M  
; APPLICANT: FRASER, PAUL E  
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED  
; TITLE OF INVENTION: TO ALZHEIMER'S DISEASE  
; NUMBER OF SEQUENCES: 183  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT  
; STREET: High Street Tower - 125 High Street  
; CITY: Boston  
; STATE: Massachusetts

COUNTRY: U.S.A.  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
FILING DATE: US/08/592,541  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Pitcher, Edmund R.  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 248-7000  
TELEFAX: (617) 248-7100  
INFORMATION FOR SEQ ID NO: 107:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 198 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-592-541-107

Query Match 1.3%; Score 41; DB 2; Length 198;  
Best Local Similarity 100.0%; Pred. No. 9.6e-09;  
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1194 ccttgccctccaaagtctggattacagcgatgagccac 1234  
|||||  
Db 12 CCTTGGCCTCCCAAGTCTGGGATTACAGGCATGAGCCAC 52

RESULT 42  
US-09-124-698-107  
; Sequence 107, Application US/09124698  
; Patent No. 6117978  
; GENERAL INFORMATION:  
; APPLICANT: ST. GEORGE-HYSLOP, PETER H  
; APPLICANT: ROMMENS, JOHANNA M  
; APPLICANT: FRASER, PAUL E  
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED  
; TITLE OF INVENTION: TO ALZHEIMER'S DISEASE  
; NUMBER OF SEQUENCES: 183  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT  
; STREET: High Street Tower - 125 High Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02110  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; FILING DATE: US/09/124,698  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/592,541  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pitcher, Edmund R.  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 107:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 198 base pairs

```

; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-124-698-107

Query Match 1.3%; Score 41; DB 3; Length 198;
Best Local Similarity 100.0%; Pred. No. 9.6e-09;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1194 ccttgccctcccaagtgtggattacaggcatgagccac 1234
|||||
DB 12 CCTTGGCCTCCCAAGTGTGGATTACAGGCATGAGCCAC 52

RESULT 43
US-09-127-480-107
; Sequence 107, Application US/09127480
; Patent No. 6194153
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 183
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: High Street Tower - 125 High Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/127,480
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,541
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 107:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-09-127-480-107

Query Match 1.3%; Score 41; DB 4; Length 198;
Best Local Similarity 100.0%; Pred. No. 9.6e-09;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1194 ccttgccctcccaagtgtggattacaggcatgagccac 1234
|||||
DB 12 CCTTGGCCTCCCAAGTGTGGATTACAGGCATGAGCCAC 52

RESULT 44
US-08-496-841C-107
; Sequence 107, Application US/08496841C
; Patent No. 6210919
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 175
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Darby & Darby, PC
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/496,841C
; FILING DATE: 28-Jun-1995
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul F. Fehlner, Ph.D.
; REGISTRATION NUMBER: 35,135
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 527-7700
; TELEFAX: (212) 753-6237
; INFORMATION FOR SEQ ID NO: 107:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 107:
US-08-496-841C-107

Query Match 1.3%; Score 41; DB 4; Length 198;
Best Local Similarity 100.0%; Pred. No. 9.6e-09;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1194 ccttgccctcccaagtgtggattacaggcatgagccac 1234
|||||
DB 12 CCTTGGCCTCCCAAGTGTGGATTACAGGCATGAGCCAC 52

RESULT 45
US-08-967-101-103
; Sequence 103, Application US/08967101
; Patent No. 5840540
; GENERAL INFORMATION:
; APPLICANT: ST. GEORGE-HYSLOP, PETER H
; APPLICANT: ROMMENS, JOHANNA M
; APPLICANT: FRASER, PAUL E
; TITLE OF INVENTION: GENETIC SEQUENCES AND PROTEINS RELATED
; TO ALZHEIMER'S DISEASE
; NUMBER OF SEQUENCES: 183
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: High Street Tower - 125 High Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/967,101
; FILING DATE: 10-NOV-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/592,541
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 386 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-967-101-103

```

```

Query Match      1.3%; Score 41; DB 2; Length 386;
Best Local Similarity 100.0%; Pred. No. 9.3e-09;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1194 cctggcctcccaagtgtggtgattacagcagcagccac 1234
      |||||||||||||||||||||||||||||||||||
Db 12 CCTGGCCTCCCAAGTGTGGGATTACAGCAGCAGCCAC 52

```

```

Search completed: September 20, 2002, 06:14:47
Job time: 11106 sec

```

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 00:17:46 ; Search time 3900.56 Seconds  
(without alignments)  
11180.103 Million cell updates/sec

Title: US-09-846-456-1  
Perfect score: 3231  
Sequence: 1 acagggcatgtggcagggtg.....gccccacatccccaccactt 3231

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 13736207 seqs, 6748477542 residues

Word size : 0

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST :\*

- 1: em\_estba:\*
- 2: em\_esthum:\*
- 3: em\_estin:\*
- 4: em\_estmu:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estro:\*
- 8: em\_hic:\*
- 9: gb\_est1:\*
- 10: gb\_est2:\*
- 11: gb\_hic:\*
- 12: gb\_gss:\*
- 13: em\_gss\_hum:\*
- 14: em\_gss\_inv:\*
- 15: em\_gss\_pln:\*
- 16: em\_gss\_vrt:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	336	10.4	763	9 AU121731	AU121731 AU121731
2	217	6.7	736	9 AU135588	AU135588 AU135588
3	154	4.8	292	10 Z44377	Z44377 HSC12B081 n
4	126	3.9	289	9 AW748338	AW748338 RC6-BT025
5	111	3.4	467	9 AA527406	AA527406 ng37c02.s
6	98	3.0	998	10 BG678861	BG678861 602624760
7	58	1.8	617	9 BE178833	BE178833 PM0-HT060
8	57	1.8	219	12 AQ538994	AQ538994 RPC1-11-3
9	57	1.8	424	9 AA652813	AA652813 ns74c05.s
10	57	1.8	674	12 AG067644	AG067644 Pan trogl
11	56	1.7	569	10 BI059823	BI059823 IL3-UT011
12	56	1.7	600	10 BI015195	BI015195 MR4-ET014
13	56	1.7	1054	10 BG745912	BG745912 602724092
14	55	1.7	384	10 BG014649	BG014649 IL5-GN023
15	55	1.7	387	10 BG007260	BG007260 IL5-GN023
16	55	1.7	444	12 AQ056873	AQ056873 CIT-HSP-2
17	54	1.7	325	9 AA010265	AA010265 zi09e05.s

18	54	1.7	405	10 BF431825	BF431825 nab50h08.
19	54	1.7	468	9 AW104031	AW104031 xg64a05.x
20	54	1.7	814	10 BG576148	BG576148 602597119
21	53	1.6	453	12 AQ663862	AQ663862 HS_2151_B
22	53	1.6	658	9 AI610607	AI610607 tpi9g09.x
23	52	1.6	233	10 BI062371	BI062371 IL3-UT011
24	52	1.6	249	10 BI710918	BI710918 id94g04.y
25	52	1.6	348	10 BI711369	BI711369 id94g04.x
26	52	1.6	391	9 AI264119	AI264119 qi08h11.x
27	52	1.6	414	9 AA481408	AA481408 zv02g09.s
28	52	1.6	416	9 AI679442	AI679442 tu63h07.x
29	52	1.6	417	9 AI337065	AI337065 qx82g11.x
30	52	1.6	417	9 AI679952	AI679952 tu67h07.x
31	52	1.6	450	10 BI062368	BI062368 IL3-UT011
32	52	1.6	455	9 AI089524	AI089524 oz22a01.x
33	52	1.6	484	9 AI002969	AI002969 au23g03.s
34	52	1.6	501	12 AQ489571	AQ489571 RPT383813
35	52	1.6	530	9 AW971724	AW971724 EST383813
36	52	1.6	545	12 AQ623696	AQ623696 HS_5319_A
37	52	1.6	583	9 AI084593	AI084593 ox63f01.s
38	52	1.6	717	12 AQ589333	AQ589333 CITBI-E1-
39	52	1.6	753	9 AV718287	AV718287 AV718287
40	51	1.6	171	12 AQ198714	AQ198714 RPT3811-59
41	51	1.6	232	9 AA662976	AA662976 ac52b01.s
42	51	1.6	360	10 H90008	H90008 yu83d03.s1
43	51	1.6	369	9 AA486970	AA486970 ab17f12.r
44	51	1.6	375	10 BF924753	BF924753 CM1-NT020
45	51	1.6	379	12 AQ280600	AQ280600 CITBI-E1-

## ALIGNMENTS

### RESULT 1

AU121731  
LOCUS AU121731 MAMMAL Homo sapiens cdna clone MAMMA1000851 5', mRNA  
DEFINITION AU121731 763 bp mRNA linear EST 19-OCT-2000  
ACCESSION AU121731  
VERSION AU121731.1 GI:10936966  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 763)  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and Isogai, T.  
HRI human cDNA project  
Unpublished (2000)  
Contact: Takao Isogai  
Genomics Laboratory  
Helix Research Institute  
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
Tel: 81-438-52-3951  
Fax: 81-438-52-3952  
Email: genomics@hri.co.jp

HRI human cDNA project; 5' & 3' end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.  
Location/Qualifiers  
1. 763  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="MAMMA1000851"  
/clone\_lib="MAMMA1"  
/tissue\_type="mammary gland"  
/note="Vector: pME18SFL3"

### FEATURES

137 a 205 c 260 g 158 t 3 others  
BASE COUNT  
ORIGIN

	Matches	217;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	2988	tctcgagcgaagtgtgaggccggacccgcagagaccgcagaccttctctccgg	2957							
Db	5	TTCGAGCGAGATGTAGTGGGCCGGACCCGCAGAGCGCAGCCCTTCTCTCCGG	64							
Qy	2958	gctcgcgcaggcgaaggccgcgcgcgcacacacagacgcggtttctcaggggcctt	3017							
Db	65	GCTCGCGCAGGCGAGGCGGGAGCTCCGGCGACCAACAGACCGCGTTCTCAGGGCGCTT	124							
Qy	3018	tgctccctgtttttccccggttctgttttctccccttctccggaagccttgtcaagg	3077							
Db	125	TGCTCCCTGTGTTTTTCCC CGGTCTCTGTGTTTCTCCCCCTCCGGAAGGCTTGTC AAGGG	184							
Qy	3078	taggaaaaagagacgcacaacaaaagtgtgaaaaag	3114							
Db	185	TAGAGAAAAGAGAGCGCAAAACACAAAAGTGTGAAAAACAG	221							

RESULT	3	
Z44377		
LOCUS	292 bp	mRNA linear EST 14-NOV-1994
DEFINITION	HSC112B081 normalized infant brain cDNA Homo sapiens CDNA clone c-12B08, mRNA sequence.	
ACCESSION	Z44377.1	GI:573506
VERSION	EST.	
KEYWORDS	human.	
SOURCE	Homo sapiens	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1 (bases 1 to 292)	
AUTHORS	Auffray,C., Behar,G., Bois,F., Bouchier,C., da Silva,C., Devignes,M.D., Duprat,S., Houlgatte,R., Jumeau,M.N., Lamy,B., Lorenzo,F., Mitchell,H., Mariage-Samson,R., Pietu,G., Pouliot,Y., Sebastiani-Kabaktchis,C. and Tessier,A.	
TITLE	IMAGE: molecular integration of the analysis of the human genome and its expression	
JOURNAL	C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)	
MEDLINE	95277534	
COMMENT	Contact: Genethon Genexpress-Genethon Genethon Centre de recherche sur le Genome Humain 1, rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE Tel: 33169472800 Fax: 33160778698 Email: genexpress@genethon.fr Single read. Genexpress_library_idt: C; Genexpress_sequence_idt: ylc-1zb08 Seq primer: (-21)M13 universal.	

```

FEATURES
source
    loc Planet: ( 219052)universals
    Location/Qualifiers
    1. .292
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="c-lzb08"
    /clone_lib="normalized infant brain cdna"
    /sex="Female"
    /tissue_type="total brain"
    /dev_stage="3 months old"
    /note="Organ: brain; Vector: lafmid BA; Site.1: HindIII;
    Site.2: NotI; sex:Female; dev_stage=3 months old;
    isolate-muscular atrophy patient; tissue_type=total brain
    ; total mRNA was oligo-(dr) primed and directionally
    cloned 5' -> 3' into the HindIII -> NotI sites of the
    lafmid BA vector. Clone library from B.Soare, Psychiatry
    Dept. Columbia University, USA.
    Bento Soares, P.N.A.S in press"
    50 a 87 c 96 g 56 t 3 others
BASE COUNT
ORIGIN

```

Query Match 4.8%; Score 154; DB 10; Length 292;  
Best Local Similarity 99.5%; Pred. No. 5.7e-52;



```

Db      1 TTTGGGCTCTCTCTCAATTATGAAGAGCAGTAGTTCCTCGGTCC 60
QY      2212 tctgaggacctgggagctcaggctggaatctccaaggcagtaggtcgc 2262
Db      61 TCTGAGGACCTGGGAGCTCAGGCTGGGATCTCCAGGAGCAGTAGGTCC 111

RESULT  6
BG678861 602624760F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4749735 5',
LOCUS    mRNA sequence.
DEFINITION
ACCESSION BG678861
VERSION   BG678861.1 GI:13910258
KEYWORDS  EST.
SOURCE    human.
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 998)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
           Email: cgaaps-r@mail.nih.gov
           Tissue Procurement: James Cleaver, M.D.
           cDNA Library Preparation: Life Technologies, Inc.
           DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA sequencing by: Incyte Genomics, Inc.
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLNL0603 row: 9 column: 16
           High quality sequence stop: 860.
           Location/Qualifiers
             1..998
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone="IMAGE:4749735"
               /clone_lib="NCI_CGAP_Skn4"
               /tissue_type="squamous cell carcinoma"
               /lab_host="DH10B (T1 phage-resistant)"
               /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
               Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
               Average insert size 1.5kb. Library constructed by Life
               Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 285 a 233 c 244 g 236 t
ORIGIN

Query Match 3.0%; Score 98; DB 10; Length 998;
Best Local Similarity 100.0%; Pred. No. 2.2e-29;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3017 ttgtctctgtttttcccccgttctgttttctcccttctcgcggaagctgtcgaagg 3076
Db 105 TTGCTCTCTGTTTTCCTCCGGTCTGTTTCTCCCTTCTCCGGAAGCTGTCTCAAGG 164

QY 3077 gtaggagaagagacgcaacacaaagtggaaaacag 3114
Db 165 GTAGGAGAAAGAGACGCAACACAAAGTGGAAAACAG 202

RESULT  7
BE178833/c BE178833
LOCUS      PMO-HT0608-170300-001-f02 HT0608 Homo sapiens cDNA, mRNA sequence.
DEFINITION
ACCESSION BE178833
VERSION   BE178833.1 GI:8657985
KEYWORDS  EST.
SOURCE    human.
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 617)
AUTHORS   Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
           Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
           Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
           Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
           , M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
           Simpson, A.J.
TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE   20202663
COMMENT    Contact: Simpson A.J.G.
           Laboratory of Cancer Genetics
           Ludwig Institute for Cancer Research
           Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
           Brazil
           Tel: +55-11-2704922
           Fax: +55-11-2707001
           Email: asimpson@ludwig.org.br
           This sequence was derived from the FAPESP/LICR Human Cancer Genome
           Project. This entry can be seen in the following URL
           (http://www.ludwig.org.br/scripts/gethtml2.pl?cl=8t2-PMO-HT0608-170
           300-001-f02&t3=2000-03-17&t4=1)
           Seq primer: puc 18 forward
           High quality sequence start: 14
           High quality sequence stop: 612.
           Location/Qualifiers
             1..617
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone_lib="HT0608"
               /dev_stage="Adult"
               /note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
               Site_2: SmaI; A mini-library was made by cloning products
               derived from ORESTES PCR (U.S. Letters Patent application
               No. 196,716 - Ludwig Institute for Cancer Research)
               profiles into the pUC 18 vector. Reverse transcription of
               tissue mRNA and cDNA amplification were performed under
               low stringency conditions."
BASE COUNT 194 a 128 c 150 g 145 t
ORIGIN

Query Match 1.8%; Score 58; DB 9; Length 617;
Best Local Similarity 100.0%; Pred. No. 5.8e-13;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1187 ctgccacattggctcccaagtgtggtattcagcagcatgagccactgcgccagc 1244
Db 247 CTGCCACCTTGGCTCCCAAGTGTGGATTACAGCATGAGCCACTGCCCCAGC 190

RESULT  8
AQ538994 AQ538994
LOCUS     RPCI-11-343L18.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-343L18
DEFINITION
ACCESSION AQ538994
VERSION   AQ538994.1 GI:4869633
KEYWORDS  GSS.
SOURCE    human.
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 219)
AUTHORS   Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and Venter
           , J.C.
TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
JOURNAL   Map Building
COMMENT    Unpublished (1997)
           Other_GSSs: RPCI-11-343L18.TV

```

```

REFERENCE 1 (bases 1 to 617)
AUTHORS   Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
           Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
           Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
           Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
           , M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
           Simpson, A.J.
TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE   20202663
COMMENT    Contact: Simpson A.J.G.
           Laboratory of Cancer Genetics
           Ludwig Institute for Cancer Research
           Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
           Brazil
           Tel: +55-11-2704922
           Fax: +55-11-2707001
           Email: asimpson@ludwig.org.br
           This sequence was derived from the FAPESP/LICR Human Cancer Genome
           Project. This entry can be seen in the following URL
           (http://www.ludwig.org.br/scripts/gethtml2.pl?cl=8t2-PMO-HT0608-170
           300-001-f02&t3=2000-03-17&t4=1)
           Seq primer: puc 18 forward
           High quality sequence start: 14
           High quality sequence stop: 612.
           Location/Qualifiers
             1..617
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone_lib="HT0608"
               /dev_stage="Adult"
               /note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
               Site_2: SmaI; A mini-library was made by cloning products
               derived from ORESTES PCR (U.S. Letters Patent application
               No. 196,716 - Ludwig Institute for Cancer Research)
               profiles into the pUC 18 vector. Reverse transcription of
               tissue mRNA and cDNA amplification were performed under
               low stringency conditions."
BASE COUNT 194 a 128 c 150 g 145 t
ORIGIN

Query Match 1.8%; Score 58; DB 9; Length 617;
Best Local Similarity 100.0%; Pred. No. 5.8e-13;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1187 ctgccacattggctcccaagtgtggtattcagcagcatgagccactgcgccagc 1244
Db 247 CTGCCACCTTGGCTCCCAAGTGTGGATTACAGCATGAGCCACTGCCCCAGC 190

RESULT  8
AQ538994 AQ538994
LOCUS     RPCI-11-343L18.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-343L18
DEFINITION
ACCESSION AQ538994
VERSION   AQ538994.1 GI:4869633
KEYWORDS  GSS.
SOURCE    human.
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 219)
AUTHORS   Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and Venter
           , J.C.
TITLE     Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
JOURNAL   Map Building
COMMENT    Unpublished (1997)
           Other_GSSs: RPCI-11-343L18.TV

```



Contact: Shaying Zhao, William Nierman, Mark Adams  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: hbe@tigr.org  
 Clones are derived from the human BAC library RPCI-11. For BAC  
 library availability, please contact Pieter de Jong  
 (pieter@dejong.med.buffalo.edu). Clones may be purchased from  
 BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from  
 Research Genet cs (info@regen.com). BAC end search page:  
 http://www.tigr.org/tldb/humgen/bac\_end\_search/bac\_end\_search.html.  
 Seq primer: SP6  
 Class: BAC ends.

## FEATURES

source

Location/Qualifiers

1. .219  
 /organism="Homo sapiens"  
 /db\_xref="GDB:7631609"  
 /db\_xref="taxon:9606"  
 /clone="RPCI-11-343L18"  
 /clone\_lib="RPCI-11"  
 /sex="Male"  
 /cell\_type="Lymphocytes"  
 /note="Vector: pBACE3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
 RPCI11 Human Male BAC Library"

BASE COUNT 41 a 63 c 35 g 80 t  
 ORIGIN

Query Match 1.8%; Score 57; DB 12; Length 219;  
 Best Local Similarity 100.0%; Pred. NO. 2.5e-12;  
 Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1025 ctcaactgaacctctgcctccgggttcaggattctctgcctcagcctcctgag 1081  
 |||||  
 Db 113 CTCAGTGTACCTCTGCCTCCGGGTTCAAGCGATTCTCTGCTCAGCCTCCTGAG 169

## RESULT 9

AA652813

LOCUS

ns74c05.s1 NCI.CGAP.Pr2 Homo sapiens cDNA clone IMAGE:1189352  
 similar to contains Alu repetitive element; , mRNA sequence.

ACCESSION

AA652813

VERSION

AA652813.1

GI:2584465

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 424)

NCI-CGAP

http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgaps@email.nih.gov

Tissue procurement: W. Marston Linehan, M.D., Rodrigo Chuaqui, M.D.

, Michael Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: David B. Krizman, Ph.D.

DNA Sequencing by: Genome Systems Inc., Greg Lennon, Ph.D.

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/brp/image/image.html

Insert Length: 560 Std Error: 0.00

Seq primer: -40m13 fwd. ET from Amersham

High quality sequence stop: 423.

Location/Qualifiers

1. .424

/organism="Homo sapiens"

/db\_xref="taxon:9606"

## FEATURES

source

/clone="IMAGE:1189352"  
 /clone\_lib="NCI.CGAP.Pr2"  
 /sex="Male"  
 /dev\_stage="45 years old"  
 /lab\_host="DH10B"  
 /note="Vector: pAMP10; Site\_1: NotI; Site\_2: EcoRI; 1st  
 strand cDNA was primed with oligo(dT)17 on 50 ng of  
 DNase-treated, total cellular RNA obtained from 5,000-10  
 ,000 microdissected preneoplastic cells  
 histologically-determined to be prostatic intraepithelial  
 neoplasia 2 (PIN2) cells. Double-stranded cDNA was  
 ligated to EcoRI adaptors, 5 cycles of PCR applied to the  
 cDNA with an adaptor-specific primer, and the resulting  
 PCR product subcloned into pAMP10 by the UDG-cloning  
 method (Life Technologies). Average insert size is 600  
 bp. NOTE: Not directionally cloned. This library was  
 constructed by David Krizman."

BASE COUNT 89 a 102 c 104 g 129 t  
 ORIGIN

Query Match 1.8%; Score 57; DB 9; Length 424;  
 Best Local Similarity 100.0%; Pred. NO. 1.8e-12;  
 Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1025 ctcaactgaacctctgcctccgggttcaggattctctgcctcagcctcctgag 1081  
 |||||  
 Db 59 CTCAGTGTACCTCTGCCTCCGGGTTCAAGCGATTCTCTGCTCAGCCTCCTGAG 115

## RESULT 10

AG067644/c

LOCUS

AG067644

DEFINITION

Pan troglodytes DNA, clone: PTB-057L02.R, genomic survey sequence.

ACCESSION

AG067644

VERSION

AG067644.1

GI:16619446

KEYWORDS

GSS; GSS (genome survey sequence).

SOURCE

Pan troglodytes male lymphoblast DNA, clone\_lib:PTB Chimpanzee Male

BAC Library clone:PTB-057L02.R.

ORGANISM

Pan troglodytes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

1 (sites)

Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,

Totoki,Y., Watanabe,H. and Sakaki,Y.

BAC end sequences of Library PTB

Unpublished

2 (bases 1 to 674)

Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,

Totoki,Y., Watanabe,H. and Sakaki,Y.

Direct Submission

Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical

and Chemical Research (RIKEN), Genomic Sciences Center (GSC);

1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

(E-mail:chimbases@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,

Tel:81-45-503-9111, Fax:81-45-503-9170)

Clones are derived from the chimpanzee BAC library PTB This BAC end

was generated during the R&amp;D process and may have higher chance of

clone tracking errors.

PRIMERS

Sequencing: M13Rev

LIBRARY

Vector : pKS145

R.Site 1 : SacI

R.Site 2 : SacI.

Location/Qualifiers

1. .674

/organism="Pan troglodytes"

/db\_xref="taxon:9598"

/clone="PTB-057L02.R"

/sex="male"

/cell\_type="lymphoblast"

/clone\_lib="PTB Chimpanzee Male BAC Library"

FEATURES

source

```

BASE COUNT      152 a      162 c      179 g      179 t      2 others
ORIGIN

Query Match
Best Local Similarity 100.0%; Pred. No. 1.4e-12; Length 674;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1025 ctcaactgtaacctgctccgcgggttcacagcattctcctgcctcagcctcctgag 1081
      |||||||
Db 219 CTCACCTGTAACTCTGCTCCCGGTTCAAGGATTCTCTCGCTCAGCTCTCTGAG 163

RESULT 11
BI059823
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 569)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil.
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ci=IL3&t2=IL3-UT0115-
080101-378-H06&t3=2001-01-08&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 568.
Location/Qualifiers
1..569
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="UT0115"
/dev_stage="Adult"
/note="Organ: uterus_tumor; Vector: puc18; Site.1: SmaI;
Site.2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT      175 a      114 c      119 g      161 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Pred. No. 4e-12; Length 569;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1190 ccacattggctcccaagtgctggattacagcagcatgagccactgcgcagct 1245
      |||||||
Db 120 CCCACCTTGGCTCCCAAGTGTCTGGGATTACAGCATGAGCCACTGCGCCGACGT 175

RESULT 12
BI015195/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 600)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil.
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ci=MR4&t2=MR4-ET0140-
190201-006-el2&t3=2001-02-19&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 72.
Location/Qualifiers
1..600
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="ET0140"
/dev_stage="Adult"
/note="Organ: lung_tumor; Vector: puc18; Site.1: SmaI;
Site.2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT      208 a      102 c      130 g      160 t
ORIGIN

Query Match
Best Local Similarity 100.0%; Pred. No. 3.9e-12; Length 600;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1189 gccacattggctcccaagtgctggattacagcagcatgagccactgcgcagc 1244
      |||||||
Db 451 GCCCACCTTGGCTCCCAAGTGTCTGGGATTACAGCATGAGCCACTGCGCCGACG 396

RESULT 13
BG745912
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
EST.
BG745912
602724092T1 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:4850375 3',
mRNA sequence.
BG745912
BG745912.1 GI:14056565
EST.

```

```

SOURCE
ORGANISM      human.
               Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1054)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/,
TITLE        National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-r@mail.nih.gov
               Tissue Procurement: Dr. Mark Watson
               cDNA Library Preparation: Ling Hong/Rubin Laboratory
               cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
               DNA sequencing by: Incyte Genomics, Inc.
               Clone distribution: MGC clone distribution information can be
               found through the I.M.A.G.E. Consortium/LLNL at:
               http://image.llnl.gov
               plate: L1CM1691 row: h column: 24
               High quality sequence stop: 841.
               Location/Qualifiers
FEATURES
source
1..1054
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /clone_lib="IMAGE:4850375"
   /clone_lib="NIH_MGC_113"
   /lab_host="DH10B (phage-resistant)"
   /note="Organ: spleen; Vector: pORF7; Site_1: XhoI; Site_2:
   EcoRI; cDNA made by oligo-dr priming. Directionally cloned
   into EcoRI/XhoI sites using the following 5' adaptor:
   GGCACGAG(G). Library constructed by Ling Hong in the
   laboratory of Gerald M. Rubin (University of California,
   Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
   Superscript II RT (Life Technologies). Note: this is a
   NIH_MGC Library."
BASE COUNT    265 a 270 c 231 g 288 t
ORIGIN

Query Match      1.7%; Score 56; DB 10; Length 1054;
Best Local Similarity 100.0%; Pred. No. 2.9e-12;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1189 gccaccttgcctcccaaaagtctggattacaggcatgagccactgcgccagc 1244
        |||||
Db 188 GCCCACCCTGGCCCTCCCAAAAGTCTGGATTACAGGATGAGCCATGCGCCCGCAGC 243

RESULT 14
BG014649
LOCUS      BG014649 384 bp mRNA linear EST 24-JAN-2001
DEFINITION IL5-GN0239-211200-342-g07 GN0239 Homo sapiens cDNA, mRNA sequence.
ACCESSION  BG014649
VERSION     BG014649.1 GI:12466063
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 384)
AUTHORS    Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
TITLE      Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
JOURNAL    Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE    20202663
COMMENT    Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL5&t2=IL5-GN0239-211200-342-g07&t3=2000-11-27&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 14
            High quality sequence stop: 371.

SOURCE
ORGANISM      human.
               Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1054)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/,
TITLE        National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-r@mail.nih.gov
               Tissue Procurement: Dr. Mark Watson
               cDNA Library Preparation: Ling Hong/Rubin Laboratory
               cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
               DNA sequencing by: Incyte Genomics, Inc.
               Clone distribution: MGC clone distribution information can be
               found through the I.M.A.G.E. Consortium/LLNL at:
               http://image.llnl.gov
               plate: L1CM1691 row: h column: 24
               High quality sequence stop: 841.
               Location/Qualifiers
FEATURES
source
1..384
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /clone_lib="GN0239"
   /dev_stage="Adult"
   /note="Organ: placenta_normal; Vector: puc18; Site_1: SmaI
   ; Site_2: SmaI; A mini-library was made by cloning
   products derived from ORESTES PCR (U.S. Letters Patent
   application No. 196,716 - Ludwig Institute for Cancer
   Research) profiles into the puc 18 vector. Reverse
   transcription of tissue mRNA and cDNA amplification were
   performed under low stringency conditions."
BASE COUNT    86 a 110 c 87 g 100 t
ORIGIN

Query Match      1.7%; Score 55; DB 10; Length 384;
Best Local Similarity 100.0%; Pred. No. 1.2e-11;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1190 ccacacttgcctcccaaaagtctggattacaggcatgagccactgcgccagc 1244
        |||||
Db 115 CCCACCTTGGCCCTCCCAAAAGTCTGGATTACAGGATGAGCCATGCGCCCGCAGC 169

RESULT 15
BG007260
LOCUS      BG007260 387 bp mRNA linear EST 24-JAN-2001
DEFINITION IL5-GN0239-271100-281-e06 GN0239 Homo sapiens cDNA, mRNA sequence.
ACCESSION  BG007260
VERSION     BG007260.1 GI:12451267
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 387)
AUTHORS    Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
TITLE      Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
JOURNAL    Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE    20202663
COMMENT    Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL5&t2=IL5-GN0239-271100-281-e06&t3=2000-11-27&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 14
            High quality sequence stop: 371.

```

100

```

Db 194 CCCACCTTGCTCCCAAGTCTGGGATTACAGCATGAGCCACTGCGCCAG 247

RESULT 18
BF431825
LOCUS
DEFINITION
nab50h08.x1 Soares NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone
IMAGE:3269606 3' similar to contains Alu repetitive element
;contains element MER22 repetitive element ;, mRNA sequence.
BF431825
ACCESSION
VERSION
KEYWORDS
SOURCE
BF431825.1 GI:11443939
EST.
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 405)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco.
High quality sequence stop: 385.
Location/Qualifiers
1. 405
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3269606"
/clone_lib="Soares.NSF_F8_9W_OT_PA_P_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Equal amounts of plasmid DNA from five normalized
libraries were mixed, and ss circles were made in vitro.
Following HAP purification, this DNA was used as tracer in
a subtractive hybridization reaction. The driver was
PCR-amplified cDNAs from pools of 5,000 clones made from
the same 5 libraries. The pools consisted of the following
libraries and cloneIDs: Soares NBHSF pool 1:
309384-310919, 323208-325895 Soares Nb2HP pool 1:
145032-147335, 147720-148103, 148872-149255, 15002 -
150407, 151176-152327 Soares Nb2Hr8-9W pool 1:
758280-760583, 772104-774407 Soares NbHPA pool 1:
304776-306311, 320136-322823, 326280-326663 Soares NBHOT
pool 1: 723720-726407, 739080-740999 Subtraction by Bento
Soares and M. Fatima Bonaldo."
BASE COUNT 83 a 113 c 84 g 125 t
ORIGIN

Query Match 1.7%; Score 54; DB 10; Length 405;
Best Local Similarity 100.0%; Pred. No. 3.1e-11;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1190 cccaccttggtcccaagtgctgggattacagcatgagccactgccccag 1243
|||||
Db 233 CCCACCTTGCTCCCAAGTCTGGGATTACAGCATGAGCCACTGCGCCAG 286

RESULT 19
AW104031
LOCUS
DEFINITION
x64a05.x1 NCI_CGAP Ov23 Homo sapiens cDNA clone IMAGE:2598512 3'
similar to contains Alu repetitive element; mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
AW104031.1 GI:6074766
EST.
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 468)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco.
High quality sequence stop: 424.
Location/Qualifiers
1. 468
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2598512"
/clone_lib="NCI_CGAP_Ov23"
/tissue_type="tumor, 5 pooled (see description)"
/lab_host="DH10B"
/note="Organ: ovary; Vector: PCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.35 kb. Tumor types include: mixed
Mullerian tumor, papillary serous, clear cell, spindle
cell. All are primary tumors, metastasis positive. Life
Technologies catalog #: 11534-013"
BASE COUNT 101 a 116 c 98 g 153 t
ORIGIN

Query Match 1.7%; Score 54; DB 9; Length 468;
Best Local Similarity 100.0%; Pred. No. 2.9e-11;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1190 cccaccttggtcccaagtgctgggattacagcatgagccactgccccag 1243
|||||
Db 250 CCCACCTTGCTCCCAAGTCTGGGATTACAGCATGAGCCACTGCGCCAG 303

RESULT 20
BG576148/c
LOCUS
DEFINITION
mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
BG576148
EST.
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 814)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10573 row: e column: 06
High quality sequence stop: 779.

```

FEATURES  
source

Location/Qualifiers  
1. .814  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4705901"  
/lab\_host="NIH\_MGC.87"  
/tissue\_type="mammary adenocarcinoma, cell line"  
/note="Organ: breast; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 1.383 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."  
240 a 208 c 197 g 169 t

BASE COUNT  
ORIGIN

Query Match 1.7%; Score 54; DB 10; Length 814;  
Best Local Similarity 100.0%; Pred. No. 2.2e-11;  
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1190 cccacctggctcccaagtgtggattacagcgcactgcgccag 1243  
|||||  
Db 172 CCCACCTGGCTCCCAAGTGTGGATTACAGGCATGAGCCACTGCCGCCAG 119

RESULT 21  
AQ663862

LOCUS  
DEFINITION HS.2151.B1.G10.T7C CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate-2151 Col-19 Row=N, DNA sequence.

ACCESSION AQ663862

VERSION AQ663862.1 GI:5171630

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 453)  
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome

JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)

MEDLINE 93980589

COMMENT Contact: Mahairas GG, Wallace JC, Hood L

High Throughput Sequencing Center

University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Clones may be purchased from Research Genetics (info@resgen.com).

BAC end web server: http://www.htsc.washington.edu

Plate: 2151 row: N column: 19

Seq primer: T7

Class: BAC ends

High quality sequence stop: 453.

Location/Qualifiers

source

1. .453  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="Plate-2151 Col-19 Row=N"  
/lab\_host="CIT Approved Human Genomic Sperm Library D"  
/sex="male"  
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in E-Coli DH10B"  
108 a 114 c 82 g 149 t

BASE COUNT  
ORIGIN

Query Match 1.6%; Score 53; DB 12; Length 453;

Best Local Similarity 100.0%; Pred. No. 7.5e-11;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1189 gccacacttgctcccaagtgtggattacagcgcactgcgcc 1241  
|||||  
Db 142 GCCACCTTGGCTCCCAAGTGTGGATTACAGGCATGAGCCACTGCCGCC 194

RESULT 22  
AI610607

LOCUS

DEFINITION

tp19g09.x1 NCI\_CGAP\_Gas4 Homo sapiens cDNA clone IMAGE:2188288 3'

similar to TR:Q99634 Q99634 RIG-G. [1] ;contains Alu repetitive

element; mRNA sequence.

ACCESSION AI610607

VERSION AI610607.1 GI:4619774

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 658)

REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

DNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www.bio.llnl.gov/bbrp/image/image.html

Insert Length: 2270 Std Error: 0.00

Seq primer: -400P from Gibco

High quality sequence stop: 372

POLYA-No.

Location/Qualifiers

source

1. .658

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:2188288"

/clone\_lib="NCI\_CGAP\_Gas4"

/tissue\_type="poorly differentiated adenocarcinoma with

signet ring cell features"

/lab\_host="DH10B"

/note="Organ: stomach; Vector: pCMV-SPORT6; Site\_1: SalI;

Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.69 kb. Life Technologies catalog #:

11549-011"

BASE COUNT 130 a 196 c 144 g 187 t 1 others

ORIGIN

Query Match

Best Local Similarity 1.6%; Score 53; DB 9; Length 658;

Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1189 gccacacttgctcccaagtgtggattacagcgcactgcgcc 1241  
|||||  
Db 243 GCCACCTTGGCTCCCAAGTGTGGATTACAGGCATGAGCCACTGCCGCC 295

RESULT 23  
BI062371

LOCUS

DEFINITION

IL3-UT0117-160301-502-E10\_1 UT0117 Homo sapiens cDNA, mRNA

sequence.

ACCESSION BI062371

VERSION BI062371.1 GI:14469898



Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
MA 02138  
Tel: 617-495-1812  
Fax: 617-495-8557  
Email: dmelton@biochem.harvard.edu

Library was constructed by Dr. J. Ferrer In vivo mass-excised to  
pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington  
University Genome Sequencing Center For information on obtaining a  
clone please contact: Dr. Hiroshi Inoue (hinoue@mol.wustl.edu)  
Possible reversed clone: similarity on wrong strand  
Seq primer: -400P from Gibco  
High quality sequence stop: 316.

## FEATURES

source

1. .348

Location/Qualifiers

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone\_lib="Human insulinoma"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: pancreas; Vector: pBluescript SK-; Site.1:

XhoI; Site.2: EcoRI; Constructed with lambda ZAPII system

(Stratagene) by Dr. J. Ferrer, in vivo mass-excised to

pBluescript SK- by Dr. H. Inoue following the Washington

University protocol

(http://genome.wustl.edu/est/lambda\_protocol.shtml).

Please contact Hiroshi Inoue, MD/PhD for further

information on this library (Metabolism Division, Permutt

Laboratory, Washington University School of Medicine, Box

8127, 660 S Euclid Ave, St. Louis, MO 63110). Note: this

is a Washington University Pancreas EST project library. "

BASE COUNT

ORIGIN

80 a 89 c 79 g 100 t

Query Match

Best Local Similarity 1.6%; Score 52; DB 10; Length 348;

Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1189

gccacccctggcctcccaagtgctggattacagcgatgacccactgcgc 1240

|||||

RESULT 26

A1264119/c

LOCUS

DEFINITION

gi08h11.x1 Soares\_NHMPu.s1 Homo sapiens cDNA clone IMAGE:1855941

3' similar to contains Alu repetitive element; mRNA sequence.

ACCESSION

A1264119

VERSION

A1264119.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 391)

NCT-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 1507 Std Error: 0.00

Seq primer: -400P from Gibco.

Location/Qualifiers

1. .391

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:1855941"

/clone\_lib="Soares\_NHMPu.s1"

/tissue\_type="Pooled human melanocyte, fetal heart, and

pregnant uterus"

/lab\_host="DH10B"

/note="Organ: mixed (see below); Vector: pT7T3D-Pac

(Pharmacia) with a modified polylinker; Site.1: Not I;

Site.2: Eco RI; Equal amounts of plasmid DNA from three

normalized libraries (melanocyte 2NBHM, pregnant uterus

NBHPU, and fetal heart NBH19W) were mixed, and ss circles

were made in vitro. Following HAP purification, this DNA

was used as tracer in a subtractive hybridization

reaction. The driver was PCR-amplified cDNAs from pools of

5,000 clones made from the same 3 libraries. The pools

consisted of I.M.A.G.E. clones 260232-265223,

340488-345479, and 484488-489479."

340488-345479, and 484488-489479."

BASE COUNT 115 a 94 c 100 g 82 t

ORIGIN

Query Match 1.6%; Score 52; DB 9; Length 391;

Best Local Similarity 100.0%; Pred. No. 2.1e-10;

Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1193

acctggcctcccaagtgctggattacagcgatgacccactgcgcgcgcgc 1244

|||||

Db 347

ACCTTGGCCTCCCAAGTGGGATTACAGGCATGAGCCACTGCGGCCACC 296

|||||

RESULT 27

AA481408

LOCUS

DEFINITION

zv02q09.s1 NCI-CGAP GCBI Homo sapiens cDNA clone IMAGE:746368 3'

similar to contains Alu repetitive element; mRNA sequence.

ACCESSION

AA481408

VERSION

AA481408.1

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 414)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,

Ph.D., Gerald Marti, M.D.

CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima

Bonaldo, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Seq primer: -41ml3 fwd. ET from Amersham

High quality sequence stop: 400.

Location/Qualifiers

1. .414

/organism="Homo sapiens"

/db\_xref="GDB:5945900"

/db\_xref="taxon:9606"

/clone="IMAGE:746368"

/clone\_lib="NCI-CGAP GCBI"

/tissue\_type="germinal center B cell"

/lab\_host="DH10B"

/note="Vector: pT7T3D-Pac (Pharmacia) with a modified

polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand

was prepared from human tonsillar cells enriched for

germinal center B cells by flow sorting (CD20+, IgD-),

provided by Dr. Louis M. Staudt (NCI), Dr. David Allman



(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was primed with a Not I - oligo(dT) primer  
[5'-GTGACCAATCTGAAGTGGAGCGCGCCCTCATTTTTTTTTTTT-3',  
]. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 106 a 112 c 86 g 110 t

Query Match 1.6%; Score 52; DB 9; Length 414;  
Best Local Similarity 100.0%; Pred. No. 2e-10;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1190 cccacctggctcccaaaagtgtggattacagcatgagcactgcgcc 1241

Db 180 CCCACCTGGCTCCCAAAAGTGTGGATTACAGGATGAGCCACGCGCC 231

## RESULT 28

AI679442 416 bp mRNA linear EST 15-DEC-1999

LOCUS tu63h07.x1 NCI\_CGAP\_Gas4 Homo sapiens cDNA clone IMAGE:2255773 3'  
DEFINITION similar to contains Alu repetitive element; contains element TAR1 repetitive element ;, mRNA sequence.

ACCESSION AI679442

VERSION AI679442.1 GI:4889624

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 416)

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Sequencing by: Greg Lennon, Ph.D.

Clone Distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 1074 Std Error: 0.00

Seq primer: -40UP from Gibco

High quality sequence stop: 400.

Location/Qualifiers

1. .416

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:2255773"

/clone\_lib="NCI\_CGAP\_Gas4"

/tissue\_type="poorly differentiated adenocarcinoma with

signed ring cell features"

/lab\_host="DH10B"

/note="Organ: stomach; Vector: pCMV-SPORT6; Site\_1: SalI;

Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.69 kb. Life Technologies catalog #:

11549-011"

BASE COUNT 83 a 106 c 102 g 125 t

ORIGIN

Query Match 1.6%; Score 52; DB 9; Length 416;

Best Local Similarity 100.0%; Pred. No. 2e-10;

Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1190 cccacctggctcccaaaagtgtggattacagcatgagcactgcgcc 1241

Db 210 CCCACCTGGCTCCCAAAAGTGTGGATTACAGGATGAGCCACGCGCC 261

## RESULT 29

AI337065

LOCUS qx82g11.x1 NCI\_CGAP\_GC6 Homo sapiens cDNA clone IMAGE:2009060 3'

DEFINITION similar to contains Alu repetitive element;,, mRNA sequence.

ACCESSION AI337065

VERSION AI337065.1 GI:4073992

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 417)

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael

R. Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima

Bonaldo, Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center

Clone Distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 528 Std Error: 0.00

Seq primer: -40UP from Gibco

High quality sequence stop: 400.

Location/Qualifiers

1. .417

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:2009060"

/clone\_lib="NCI\_CGAP\_GC6"

/tissue\_type="pooled germ cell tumors"

/lab\_host="DH10B"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site\_1: Not I; Site\_2: Eco RI; Plasmid DNA

from the normalized library NCI\_CGAP\_GC4 was prepared, and

ss circles were made in vitro. Following HAP purification,

this DNA was used as tracer in a subtractive hybridization

reaction. The driver was PCR-amplified cDNAs from a pool

of 5,000 clones made from the same library (clonids

1257096-1258631, 1469064-1470983, and 1475592-1476743).

Subtraction by Bento Soares and M. Fatima Bonaldo.

BASE COUNT 101 a 97 c 94 g 125 t

ORIGIN

Query Match 1.6%; Score 52; DB 9; Length 417;

Best Local Similarity 100.0%; Pred. No. 2e-10;

Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1189 gccacacctggctcccaaaagtgtggattacagcatgagcactgcgcc 1240

Db 323 GCCCACCTGGCTCCCAAAAGTGTGGATTACAGGATGAGCCACGCGCC 374

## RESULT 30

AI679952

LOCUS tu67h07.x1 NCI\_CGAP\_Gas4 Homo sapiens cDNA clone IMAGE:2256157 3'

DEFINITION similar to contains Alu repetitive element;,, mRNA sequence.

ACCESSION AI679952

VERSION AI679952.1 GI:4890134

EST 26-MAY-1999

```

KEYWORDS      EST.
SOURCE        human.
ORGANISM      Homo sapiens
REFERENCE     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE        1 (bases 1 to 417)
JOURNAL      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
COMMENT      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
              Tumor Gene Index
              Unpublished (1997)
              Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-r@mail.nih.gov
              Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
              Emmert-Buck, M.D., Ph.D.
              CDNA Library Preparation: Life Technologies, Inc.
              CDNA Library Arrayed by: Greg Lennon, Ph.D.
              DNA Sequencing by: Washington University Genome Sequencing Center
              Clone distribution: NCI-CGAP clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              www-bio.lnlni.gov/bbrp/image/image.html
              Seq primer: -40UP from Gibco
              High quality sequence stop: 409.
FEATURES      Location/Qualifiers
              1..417
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone_lib="NCI-CGAP_Gas4"
              /tissue_type="poorly differentiated adenocarcinoma with
              signet ring cell features"
              /lab_host="DH10B"
              /note="Organ: stomach; Vector: pCMV-SPORT6; Site_1: SalI;
              Site_2: NotI; Cloned unidirectionally. Primer: Oligo dr.
              Average insert size 1.69 kb. Life Technologies catalog #:
              11549-011"
BASE COUNT    83 a 106 c 102 g 126 t
ORIGIN

Query Match      1.6%; Score 52; DB 9; Length 417;
Best Local Similarity 100.0%; Pred. No. 2e-10;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1190 cccacttgccctccaaagtctggattacagcagcatgacccactgcgcc 1241
|||||
Db 210 CCCACTTGGCTCCCAAGTCTGGGATTACAGGCATGAGCCACTGCGCC 261

RESULT 31
BI062368      450 bp      mRNA      linear      EST 15-JUN-2001
LOCUS      IL3-UT0117-160301-502-G10 UT0117 Homo sapiens cDNA, mRNA sequence.
DEFINITION
ACCESSION      BI062368
VERSION      BI062368.1 GI:14469895
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 450)
AUTHORS      Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
              Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
              Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
              Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
              M.J., Soares,F., Brentani,R., Reis,L.F., de Souza,S.J. and
              Simpson,A.J.
              Shotgun sequencing of the human transcriptome with ORF expressed
              sequence tags
              Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
              20202663
              Contact: Simpson A.J.G.
              Laboratory of Cancer Genetics
              Ludwig Institute for Cancer Research

us-09-846-456-1.oli.rst

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL3&t2=IL3-UT0117-
160301-502-G10&t3=2001-03-16&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 450.
FEATURES      Location/Qualifiers
              1..450
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone_lib="UT0117"
              /dev_stage="Adult"
              /note="Organ: uterus_tumor; Vector: puc18; Site_1: SmaI;
              Site_2: SmaI; A mini-library was made by cloning products
              derived from ORESTES PCR (U.S. Letters Patent application
              No. 196,716 - Ludwig Institute for Cancer Research)
              profiles into the pUC 18 vector. Reverse transcription of
              tissue mRNA and cDNA amplification were performed under
              low stringency conditions."
              low stringency conditions."
BASE COUNT    111 a 122 c 97 g 120 t
ORIGIN

Query Match      1.6%; Score 52; DB 10; Length 450;
Best Local Similarity 100.0%; Pred. No. 1.9e-10;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1189 gccacttgccctccaaagtctggattacagcagcatgacccactgcgcc 1240
|||||
Db 157 GCCACCTTGGCTCCCAAGTCTGGGATTACAGGCATGAGCCACTGCGCC 208

RESULT 32
AI089524      455 bp      mRNA      linear      EST 01-OCT-1998
LOCUS      OZ22a01.x1 Soares_total_fetus_Nb2HF8_9w Homo sapiens cDNA clone
DEFINITION      IMAGE:1676040 3' similar to contains Alu repetitive element;; mRNA
              sequence.
ACCESSION      AI089524
VERSION      AI089524.1 GI:3428583
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 455)
AUTHORS      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
              National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
              Tumor Gene Index
              Unpublished (1997)
              Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-r@mail.nih.gov
              This clone is available royalty-free through LLNL; contact the
              IMAGE Consortium (info@image.llnl.gov) for further information.
              Insert Length: 761 Std Error: 0.00
              Seq primer: -40ml3 fwd. ET from Amersham
              High quality sequence stop: 444.
              Location/Qualifiers
              1..455
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /clone="IMAGE:1676040"
              /clone_lib="Soares_total_fetus_Nb2HF8_9w"
              /dev_stage="8-9 weeks"
              /lab_host="DH10B"
              /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
              polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
              was prepared from mRNA obtained from pooled 8-9 week

```

(total) fetus material with a Not I - oligo(dT) primer [5' TGTTCAATCTCAAGTGGGAGCGCCGCTTAATTTTTTTTTTTTTTTT 3']. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 136 a 105 c 111 g 103 t

Query Match 1.6%; Score 52; DB 9; Length 455;  
Best Local Similarity 100.0%; Pred. No. 1.9e-10; Indels 0; Gaps 0;  
Matches 52; Conservative 0; Mismatches 0;

Qy 1193 acctggcctcccaagtgtgggtattacagcagcagcagcagcagc 1244  
|||||  
Db 359 ACCTTGGCTCCCAAGTGTGGGATTACAGGCATGAGCCACTGCGCCAGC 308

RESULT 33  
LOCUS AI002969 484 bp mRNA linear EST 09-JUN-1998  
DEFINITION an23g03.s1 Gessler Wilms tumor Homo sapiens cDNA clone  
IMAGE:1699540 3' similar to contains Alu repetitive element  
; contains element MER22 repetitive element ;, mRNA sequence.

ACCESSION AI002969  
VERSION AI002969.1 GI:3203383  
KEYWORDS EST.  
SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 484)  
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,  
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin  
J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,  
White, Y., Wylie, T., Waterston, R. and Wilson, R.

TITLE WashU-NCI human EST Project  
JOURNAL Unpublished (1997)  
COMMENT Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810

Email: est@watson.wustl.edu  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: -40ml3 fwd. ET from Amersham  
High quality sequence stop: 448.

FEATURES  
source  
1..484  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1699540"  
/clone\_lib="Gessler Wilms tumor"  
/sex="pooled (6)"  
/lab\_host="DH10B"  
/note="Vector: pSPORT1; Site\_1: SalI; Site\_2: NotI; RNA  
was prepared from a pool of 6 anonymous Wilms' tumor RNAs.  
RNA was prepared by acid-phenol, followed by one round of  
oligo dT selection. cDNA library preparation was with  
the BRL/Life tech. Superscript Plasmid system. An  
oligo-dT NotI primer for first strand synthesis generated  
gcggccgccctn at the 3' end of the clones. A 5' SalI  
adaptor was used with sequence 5'-gtcaccacgcgtcg-3'.  
Resulting cDNAs were size selected (average size 2 kb),  
NotI digested, and ligated into NotI/SalI-cut pSPORT1.  
Library was constructed by Dr. Manfred Gessler."

BASE COUNT 126 a 131 c 103 g 124 t

Query Match 1.6%; Score 52; DB 9; Length 484;  
Best Local Similarity 100.0%; Pred. No. 1.9e-10; Indels 0; Gaps 0;  
Matches 52; Conservative 0; Mismatches 0;

Qy 1190 ccaccttgctcccaagtgtgggtattacagcagcagcagcagcagc 1241  
|||||  
Db 227 CCCACTTGGCTCCCAAGTGTGGGATTACAGGCATGAGCCACTGCGCC 278

RESULT 34  
LOCUS AQ489571 501 bp DNA linear GSS 24-APR-1999  
DEFINITION RPCI-11-230P5.TV RPCI-11 Homo sapiens genomic clone RPCI-11-230P5,  
DNA sequence.

ACCESSION AQ489571  
VERSION AQ489571.1 GI:4675445  
KEYWORDS GSS.  
SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 501)  
AUTHORS Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and Venter  
J.C.

TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready  
Map Building  
JOURNAL Unpublished (1997)

COMMENT Other\_GSSs: RPCI-11-230P5.TJ  
Contact: Shaying Zhao, William Nierman, Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: hbs@tigr.org

Clones are derived from the human BAC library RPCI-11. For BAC  
library availability, please contact Pieter de Jong  
(pieter@jong.med.buffalo.edu). Clones may be purchased from  
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from  
Research Genet cs (info@resgen.com). BAC end search page:  
http://www.tigr.org/tdb/humgen/bac\_end\_search/bac\_end\_search.html.  
Seq primer: T7  
Class: BAC ends.

FEATURES  
source  
1..501  
/organism="Homo sapiens"  
/db\_xref="GBB:7588300"  
/db\_xref="taxon:9606"  
/clone="RPCI-11-230P5"  
/clone\_lib="RPCI-11"  
/sex="Male"  
/cell\_type="Lymphocytes"  
/note="Vector: pBACe3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
RPCI11 Human Male BAC Library"

BASE COUNT 106 a 123 c 107 g 165 t

Query Match 1.6%; Score 52; DB 12; Length 501;  
Best Local Similarity 100.0%; Pred. No. 1.8e-10; Indels 0; Gaps 0;  
Matches 52; Conservative 0; Mismatches 0;

Qy 1193 acctggcctcccaagtgtgggtattacagcagcagcagcagcagc 1244  
|||||  
Db 395 ACCTTGGCTCCCAAGTGTGGGATTACAGGCATGAGCCACTGCGCCAGC 446

RESULT 35  
LOCUS AW971724 530 bp mRNA linear EST 01-JUN-2000  
DEFINITION EST383813 MAGE resequences, MAGL Homo sapiens cDNA, mRNA sequence.  
ACCESSION AW971724  
VERSION AW971724.1 GI:8161570

```

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 530)
AUTHORS Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C., Holt,
I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and
Quackenbush, J.
TITLE Assessment of gene expression patterns in a model of colon tumor
metastasis using a 19,200 element cDNA microarray
JOURNAL Unpublished (2000)
COMMENT Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: johnq@tigr.org
Plate: 294
Seq primer: Forward.
Location/Qualifiers
1..530
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="IMAGE resequencing, MAGL"
/note="Vector: pBluescriptSKm"

BASE COUNT 135 a 139 c 111 g 145 t
ORIGIN

Query Match 1.6%; Score 52; DB 9; Length 530;
Best Local Similarity 100.0%; Pred. No. 1.8e-10;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1190 cccacgtggctcccaagtgctgggattacagcagcatgagccactgcgcctc 1241
|||||
Db 227 CCCACGTTGGCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACTGGCGCC 278

RESULT 36
A0623696
LOCUS HS_5319_A2_C10_T7A RPCI-11 Human Male BAC Library Homo sapiens
DEFINITION genomic clone Plate=895 Col=20 Row=E, DNA sequence.
ACCESSION A0623696
VERSION A0623696.1 GI:5086011
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 545)
AUTHORS Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and
Hood, L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U. S. A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering.bac.htm)
or from Resear h Genetics (info@resgen.com). BAC end Web Server:
http://www.htsc.washington.edu
Plate: 895 row: E column: 20

Seq primer: T7
Class: BAC ends
High quality sequence stop: 545.
Location/Qualifiers
1..545
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=895 Col=20 Row=E"
/clone_lib="RPCI-11 Human Male BAC Library"
/sex="male"
/note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI Methylase. Size selected DNA was cloned into the
pBACE3.6 vector at EcoRI sites"

BASE COUNT 117 a 139 c 139 g 144 t 6 others
ORIGIN

Query Match 1.6%; Score 52; DB 12; Length 545;
Best Local Similarity 100.0%; Pred. No. 1.7e-10;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1194 ccttgctcccaagtgctgggattacagcagcatgagccactgcgcctc 1245
|||||
Db 349 CTTGGCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACTGGCCCACT 400

RESULT 37
A1084593/c
LOCUS A1084593 583 bp mRNA linear EST 28-AUG-1998
DEFINITION ox63f01.s1 Soares_NHMPu_S1 Homo sapiens cDNA clone IMAGE:1661017
3' similar to contains Alu repetitive element;; mRNA sequence.
ACCESSION A1084593
VERSION A1084593.1 GI:3423016
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 583)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 470 Std Error: 0.00
Seq primer: -40ml3 fwd. Et from Amersham
High quality sequence stop: 411.
Location/Qualifiers
1..583
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1661017"
/clone_lib="Soares_NHMPu_S1"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below); Vector: pTT73D-Pac
(Pharmacia) with a modified polylinker; Site_1: Not I;
Site_2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NBHM, pregnant uterus
NBHPU, and fetal heart NBHH19W) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."

BASE COUNT 165 a 138 c 139 g 140 t 1 others

```

```

ORIGIN
Query Match      1.6%; Score 52; DB 9; Length 583;
Best Local Similarity 100.0%; Pred. No. 1.7e-10;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1193 accttgccctccaaagtctggattacaggatgagccactgcgccagc 1244
|||||
Db 346 ACCTTGCCCTCCCAAAAGTCTGGGATTACAGGCATGAGCCACTGCGCCCGCAGC 295

RESULT 38
AQ589333/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and
Venter,J.C.
TITLE
Use of BAC End Sequences from CalTech Libraries for Sequence-Ready
Map Building
JOURNAL
COMMENT
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.
FEATURES
Location/Qualifiers
1..717
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2645K6"
/clone_lib="CITBI-EI"
/sex="male"
/cell_type="sperm"
/note="Vector: pBeloBAC11; Site_1: EcoRI; Site_2: EcoRI;
Caltech Human BAC Library D"
BASE COUNT 222 a 175 c 140 g 180 t
ORIGIN

Query Match      1.6%; Score 52; DB 12; Length 717;
Best Local Similarity 100.0%; Pred. No. 1.5e-10;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1025 ctactgtaacctgctccgggttcacggattctcctgcctcagccctc 1076
|||||
Db 464 CTCACGTGAACCTTCGCTCCGGGTTCAAGGATTCTCCTGCTCAGCCCTC 413

RESULT 39
AV718287/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
human.
AV718287
AV718287 FHTB Homo sapiens cDNA clone FHTBABH02 5', mRNA sequence.
AV718287
AV718287.1 GI:10815439
EST.
human.

```

---

```

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
Xiao,H., Peng,Y., Song,H., Gu,Y., Yang,Y., Gao,G., Xu,X., Li,N., Gu
Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., Gu
W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Chen,Z. and Han
Z.
TITLE
Homo sapiens cDNA FHTB clones
JOURNAL
COMMENT
Unpublished (2000)
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzgchgc.sh.cn
This clone is available at CHGC in Shanghai.
FEATURES
Location/Qualifiers
1..753
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="FHTBABH02"
/clone_lib="FHTB"
/tissue_type="hypothalamus"
/dev_stage="Fetal"
/lab_host="BM25.8"
/note="Vector: pTriplex2; Site_1: sfIIA; Site_2: sfIIB1"
BASE COUNT 214 a 186 c 182 g 171 t
ORIGIN

Query Match      1.6%; Score 52; DB 9; Length 753;
Best Local Similarity 100.0%; Pred. No. 1.5e-10;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1187 ctgccaccttgccctccaaagtctggattacaggatgagccactgcg 1238
|||||
Db 52 CTGCCACCTTGCCCTCCCAAAAGTCTGGGATTACAGGCATGAGCCACTGCG 1

RESULT 40
AQ198714/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
human.
AQ198714
AQ198714.1 GI:3610913
GSS.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
TITLE
Use of human BAC End Sequences for Sequence-Ready Map Building
JOURNAL
COMMENT
Unpublished (1998)
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Class: BAC ends.
FEATURES
Location/Qualifiers

```

```

source      1. 171
/organism="Homo sapiens"
/db_xref="GDB:7522510"
/db_xref="taxon:9606"
/clone="RPC1-11-59723"
/clone_lib="RPC1-11"
/sex="Male"
/cell_type="Lymphocytes"
/notes="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
RPC111 Human Male BAC Library"
BASE COUNT 43 a 32 c 44 g 52 t
ORIGIN

Query Match      1.6%; Score 51; DB 12; Length 171;
Best Local Similarity 100.0%; Pred. No. 8e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1190 cccacctgctcccaagtgctgggattacaggcatgagccactgcgcc 1240
|||||
Db 100 CCCACCTTGCTCCCAAGTGTGGGATTACAGGCATGAGCCACTGCGCC 50

RESULT 41
AA662976      232 bp mRNA linear EST 12-NOV-1997
LOCUS ac52b01.s1 Stratagene fetal retina 937202 Homo sapiens cDNA clone
DEFINITION IMAGE:866185 3' similar to contains Alu repetitive element; mRNA
sequence.
ACCESSION AA662976
VERSION AA662976
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 232)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
White,Y., Wylie,T., Waterston,R. and Wilson,R.
WASHU-NCI human EST Project
Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stop: 216.
FEATURES
source
1. 232
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:866185"
/clone_lib="Stratagene fetal retina 937202"
/sex="mixed"
/lab_host="SOLR (kanamycin resistant)"
/notes="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; Cloned unidirectionally. Primer: Oligo dT. Pooled
retinal tissue. Average insert size: 1.0 kb; Uni-ZAP XR
vector; -5' adaptor sequence: 5' GAATTCGGCAGCAG 3' -3'
adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"
BASE COUNT 48 a 51 c 55 g 78 t
ORIGIN

Query Match      1.6%; Score 51; DB 9; Length 232;
Best Local Similarity 100.0%; Pred. No. 6.9e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1031 gtaacctgctcccggttcaagcgattctctgctcgaagcctcctgag 1081
|||||
Db 115 GTAACCTCTGCTCCCGGTTCAAGCGATTCTCTGCTCAGCCTCCTGAG 165

RESULT 42
H90008/c      360 bp mRNA linear EST 28-NOV-1995
LOCUS YU83d03.s1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
DEFINITION IMAGE:240389 3' similar to contains Alu repetitive element; contains
MER26 repetitive element ;, mRNA sequence.
ACCESSION H90008
VERSION H90008.1 GI:1080438
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 360)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
Rifkin,L., Rohlfing,I., Soares,M., Tan,F., Trevasakis,E., Waterston
R., Williamson,A., Wohlmann,P. and Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 320
Insert Size: 1095
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1095 Std Error: 0.00
Seq primer: Promega -21ml3
High quality sequence stop: 320.
FEATURES
source
1. 360
/organism="Homo sapiens"
/db_xref="GDB:3789342"
/db_xref="taxon:9606"
/clone="IMAGE:240389"
/clone_lib="Soares fetal liver spleen INFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/notes="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5' AACTGGAGAGATTAATTAAGATCTTTTTTTTTTTTTTTT 3',
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Patima Bonaldo."
BASE COUNT 108 a 70 c 81 g 98 t 3 others
ORIGIN

Query Match      1.6%; Score 51; DB 10; Length 360;
Best Local Similarity 100.0%; Pred. No. 5.5e-10;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1194 ccttgccctcccaagtgctgggattacaggcatgagccactgcgccagc 1244
|||||
Db 174 CCTTGCCCTCCCAAGTGTGGGATTACAGGCATGAGCCACTGCGCCAGC 124

RESULT 43

```

TITLE	<p>,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.  Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)</p>									
JOURNAL	20202663									
MEDLINE	Contact: Simpson A.J.G. Laboratory of Cancer Genetics Ludwig Institute for Cancer Research Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil Tel: +55-11-2704922 Fax: +55-11-2707001									
COMMENT	Email: asimpson@ludwig.org.br This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=CM1&t2=CM1-NT0209-281100-609-cl0&t3=2000-11-28&t4=1) Seq primer: puc 18 forward High quality sequence start: 15 High quality sequence stop: 374. Location/Qualifiers 1. .375 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_lib="NT0209" /dev_stage="Adult" /note="Organ: nervous_tumor; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions." 99 a 94 c 113 g 69 t									
BASE COORDINATE	99 a 94 c 113 g 69 t									
ORIGIN	Query Match 1.6% Score 51; DB 10; Length 375; Best Local Similarity 100.0%; Pred. No. 5.4e-10; Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1190	ccacactggcctcccaagtgcctgggtacagggcatgacactgcgc 1240       								
Db	165	CCACCTTGGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACTGCGCC 115								
RESULT	45									
ACCESSION	AQ280600	379 bp DNA linear GSS 22-NOV-1998								
LOCUS	CITBI-E1-2516H22.TR	CITBI-E1 Homo sapiens genomic clone 2516H22, DNA sequence.								
DEFINITION										
VERSION	AQ280600.1	GI:3906419								
KEYWORDS	GSS.									
SOURCE	human.									
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 379) Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and Venter,J.C. Use of a random human BAC End Sequence Database for Sequence-Ready Map Building Unpublished (1998) Other GSSs: CITBI-E1-2516H22.TF Contact: Mark Adams Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0200 Fax: 301 838 0208									
REFERENCE										
AUTHORS										
TITLE										
JOURNAL										
COMMENT										

Email: mdadamsetigr.org  
 Clones are available from Research Genetics (info@resgen.com). BAC  
 end search page:  
[http://www.tigr.org/tdb/hungen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html).  
 Seq primer: M13 Reverse  
 Class: BAC ends.

# FEATURES

## Source

Location/Qualifiers  
 1..379  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="2516H22"  
 /clone\_lib="CITBI-E1"  
 /sex="male"  
 /cell\_type="sperm"  
 /note="Vector: pBelOAC11; Site\_1: EcoRI; Site\_2: EcoRI;  
 Caltech Human BAC Library D"

BASE COUNT  
 ORIGIN

85 a 98 c 111 g 85 t

Query Match 1.6%; Score 51; DB 12; Length 379;  
 Best Local Similarity 100.0%; Pred. No. 5.4e-10;  
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1194 ccttggcctcccaaaagtctggtattacagcagcactgcccagc 1244  
 ||||||||||||||||||||||||||||||||||||||||||||  
 DB 174 CCTTGGCCTCCCAAGTCTGGGATTACAGGCGTGGCCCGCCAGC 224

Search completed: September 20, 2002, 04:07:18  
 Job time: 13772 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 06:25:43 ; Search time 5250.46 Seconds  
(without alignments)  
1422.880 Million cell updates/sec

Title: US-09-846-456-2  
Perfect score: 357  
Sequence: 1 tggaggctcagctgaggg.....gagggaaggaagctgtgttg 357

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10463268293 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

1: gb_ba.*	AX351030	Sequence	6	AX351030	100.0	357	6	AX351030	AX351030	Sequence	357 bp	DNA	linear	PAT 06-FEB-2002
2: gb_htg.*	AF275948	Homo sapi	255	71.4	149034	9	AF275948	AF275948	AF275948	Homo sapi	Sequence			
3: gb_in.*	AX351033	Sequence	159	44.5	159	9	AX351033	AX351033	AX351033	Sequence				
4: gb_om.*	HS252277	Homo sapi	159	44.5	480	9	HS252277	HS252277	HS252277	Homo sapi	Sequence			
5: gb_ov.*	AL353685	Human DNA	159	44.5	129608	9	AL353685	AL353685	AL353685	Human DNA	Sequence			
6: gb_pat.*	AC012230	Homo sapi	159	44.5	175064	2	AC012230	AC012230	AC012230	Homo sapi	Sequence			
7: gb_ph.*	AC092589	Sequence	159	44.5	183999	6	AC092589	AC092589	AC092589	Sequence				
8: gb_pr.*	AF287262	Homo sapi	159	44.5	201144	9	AF287262	AF287262	AF287262	Homo sapi	Sequence			
9: gb_ro.*	AC021345	Homo sapi	116	32.5	90698	2	AC021345	AC021345	AC021345	Homo sapi	Sequence			
10: gb_sy.*	AX351035	Sequence	99	27.7	99	6	AX351035	AX351035	AX351035	Sequence				
11: gb_un.*	AF258624	Homo sapi	81	22.7	200	9	AF258624	AF258624	AF258624	Homo sapi	Sequence			
12: gb_vi.*	AX060713	Sequence	77	21.6	10442	6	AX060713	AX060713	AX060713	Sequence				
13: gb_vl.*	AF060892	Sequence	77	21.6	10442	6	AF060892	AF060892	AF060892	Sequence				
14: gb_ba.*	AF285167	Homo sapi	77	21.6	10442	6	AF285167	AF285167	AF285167	Homo sapi	Sequence			
15: em_ba.*	AX060721	Sequence	77	21.6	10474	6	AX060721	AX060721	AX060721	Sequence				
16: em_fun.*	AX060898	Sequence	77	21.6	10474	6	AX060898	AX060898	AX060898	Sequence				
17: em_hum.*	AB037924	Homo sapi	77	21.6	10474	6	AB037924	AB037924	AB037924	Homo sapi	Sequence			
18: em_in.*	AX127764	Sequence	60	16.8	298	9	AX127764	AX127764	AX127764	Sequence				
19: em_mu.*	AX139751	Sequence	60	16.8	446	6	AX139751	AX139751	AX139751	Sequence				
20: em_om.*	AF258627	Homo sapi	60	16.8	697	9	AF258627	AF258627	AF258627	Homo sapi	Sequence			
21: em_or.*	AB055982	Homo sapi	60	16.8	6786	9	AB055982	AB055982	AB055982	Homo sapi	Sequence			
22: em_ov.*	AX253452	Sequence	60	16.8	7260	6	AX253452	AX253452	AX253452	Sequence				
23: em_ph.*	AX092594	Sequence	60	16.8	7860	6	AX092594	AX092594	AX092594	Sequence				
24: gb_pr.*	AX135712	Sequence	60	16.8	7862	6	AX135712	AX135712	AX135712	Sequence				
25: gb_ro.*	AX127830	Sequence	60	16.8	9741	6	AX127830	AX127830	AX127830	Sequence				
26: gb_sy.*	AX139817	Sequence	60	16.8	9741	6	AX139817	AX139817	AX139817	Sequence				
27: gb_un.*	AX351038	Sequence	60	16.8	9741	6	AX351038	AX351038	AX351038	Sequence				
28: gb_vl.*	AX127831	Sequence	60	16.8	9854	6	AX127831	AX127831	AX127831	Sequence				
29: gb_vi.*	AK024328	Homo sapi	60	16.8	9854	6	AK024328	AK024328	AK024328	Homo sapi	Sequence			
30: em_ba.*	AX092843	Sequence	51	14.3	1556	9	AX092843	AX092843	AX092843	Sequence				
31: em_fun.*	AX092841	Sequence	35	9.8	37	6	AX092841	AX092841	AX092841	Sequence				
32: em_hum.*	AC091456	Mus muscu	35	9.8	38	6	AC091456	AC091456	AC091456	Mus muscu	Sequence			
33: em_in.*	AF287263	Mus muscu	35	9.8	38	6	AF287263	AF287263	AF287263	Mus muscu	Sequence			
34: em_mu.*	AX239607	Sequence	37	7.6	207659	2	AX239607	AX239607	AX239607	Sequence				
35: em_om.*	AC006326	Homo sapi	37	7.6	278572	10	AC006326	AC006326	AC006326	Homo sapi	Sequence			
36: gb_ph.*	AL583882	Human DNA	38	6.7	12425	6	AL583882	AL583882	AL583882	Human DNA	Sequence			
37: gb_ro.*	AF179296	Homo sapi	39	6.7	18427	9	AF179296	AF179296	AF179296	Homo sapi	Sequence			
38: gb_sy.*	AC036144	Homo sapi	40	6.7	90582	9	AC036144	AC036144	AC036144	Homo sapi	Sequence			
39: gb_un.*	AL158043	Human DNA	42	6.7	123399	2	AL158043	AL158043	AL158043	Human DNA	Sequence			
40: gb_vl.*	AL163512	Mouse DNA	42	6.7	127270	9	AL163512	AL163512	AL163512	Mouse DNA	Sequence			
41: gb_vi.*	AL353893	Homo sapi	44	6.7	132741	9	AL353893	AL353893	AL353893	Homo sapi	Sequence			
42: em_ba.*			44	6.7	134334	10								
43: em_fun.*			44	6.7	137072	2								
44: em_hum.*			45	6.7	137072	2								
45: em_in.*														

ALIGNMENTS

RESULT 1

LOCUS AX351030  
DEFINITION Sequence 2 from Patent WO0183746.  
ACCESSION AX351030  
VERSION AX351030.1 GI:18616386  
KEYWORDS  
SOURCE human.

ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (sites)  
Rosier-Montus, M.F., Prades, C., Lemoine, C., Naudin, L., Deneffe, P.,  
Brewer, B., Duverger, N., Remaley, A. and Santamarina-Fojo, S.  
Regulatory nucleic acid sequences of the abcl gene  
Patent: WO 0183746-A 2 08-NOV-2001;  
Aventis Pharma S.A. (FR)  
Location/Qualifiers  
source  
1. .357  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

BASE COUNT

ORIGIN

Result No. Score Match Length DB ID Description

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES







repeat\_region 3237. .3415  
 /note="LIME repeat: matches 5696. .5821 of consensus"  
 repeat\_region 6522. .6818  
 /note="AluSq repeat: matches 1. .295 of consensus"  
 repeat\_region 7282. .7415  
 /note="LIMB8 repeat: matches 6040. .6173 of consensus"  
 repeat\_region 8145. .8434  
 /note="AluSc repeat: matches 1. .298 of consensus"  
 repeat\_region 12145. .12713  
 /note="L2 repeat: matches 1363. .1940 of consensus"  
 repeat\_region 13890. .13969  
 /note="L2 repeat: matches 2611. .2701 of consensus"  
 repeat\_region 15380. .15411  
 /note="16 copies 2 mer ac 87% conserved"  
 repeat\_region 16105. .16144  
 /note="10 copies 4 mer caca 100% conserved"  
 repeat\_region 16868. .17049  
 /note="MIR repeat: matches 64. .246 of consensus"  
 repeat\_region 17941. .18229  
 /note="AluSq repeat: matches 1. .287 of consensus"  
 repeat\_region 18259. .18553  
 /note="AluSq repeat: matches 1. .293 of consensus"  
 repeat\_region 20310. .20616  
 /note="AluSq repeat: matches 1. .308 of consensus"  
 repeat\_region 20957. .21107  
 /note="MIR repeat: matches 49. .212 of consensus"  
 repeat\_region 21783. .22078  
 /note="AluSc repeat: matches 7. .302 of consensus"  
 repeat\_region 22320. .22439  
 /note="MIR repeat: matches 10. .146 of consensus"  
 repeat\_region 22533. .22839  
 /note="AluSc repeat: matches 1. .307 of consensus"  
 repeat\_region 23427. .23945  
 /note="L2 repeat: matches 2137. .2750 of consensus"  
 repeat\_region 24245. .24544  
 /note="AluSq repeat: matches 2. .302 of consensus"  
 repeat\_region 24556. .24587  
 /note="8 copies 4 mer acac 96% conserved"  
 repeat\_region 26504. .26561  
 /note="29 copies 2 mer ta 69% conserved"  
 repeat\_region 26849. .26892  
 /note="11 copies 4 mer tata 81% conserved"  
 repeat\_region 27998. .28103  
 /note="MIR repeat: matches 28. .145 of consensus"  
 repeat\_region 28515. .28626  
 /note="MIR repeat: matches 17. .129 of consensus"  
 repeat\_region 28986. .29213  
 /note="MIR repeat: matches 2. .245 of consensus"  
 repeat\_region 30121. .30422  
 /note="AluSq repeat: matches 1. .302 of consensus"  
 repeat\_region 31424. .31734  
 /note="AluSp repeat: matches 1. .308 of consensus"  
 repeat\_region 31987. .32116  
 /note="MER5A repeat: matches 60. .188 of consensus"  
 repeat\_region 32450. .32536  
 /note="MER5A repeat: matches 14. .106 of consensus"  
 repeat\_region 33293. .33380  
 /note="22 copies 4 mer atgg 79% conserved"  
 repeat\_region 34435. .34588  
 /note="MER5A repeat: matches 13. .189 of consensus"  
 repeat\_region 34729. .34873  
 /note="MIR repeat: matches 49. .198 of consensus"  
 repeat\_region 35802. .35951  
 /note="MIR repeat: matches 77. .248 of consensus"  
 repeat\_region 37183. .37260  
 /note="2 copies 39 mer 92% conserved"  
 repeat\_region 37673. .37980  
 /note="AluSc repeat: matches 1. .308 of consensus"  
 repeat\_region 39674. .40243  
 /note="LIMD2 repeat: matches 5774. .6331 of consensus"  
 repeat\_region 40256. .40534  
 /note="L2 repeat: matches 2256. .2533 of consensus"  
 repeat\_region 41476. .41615

repeat\_region 42010. .42194  
 /note="MIR repeat: matches 30. .185 of consensus"  
 repeat\_region 44809. .45239  
 /note="LTR16A repeat: matches 6. .450 of consensus"  
 repeat\_region 45359. .45486  
 /note="MIR repeat: matches 1. .139 of consensus"  
 repeat\_region 46654. .46693  
 /note="10 copies 4 mer acac 97% conserved"  
 repeat\_region 47118. .47429  
 /note="AluSc repeat: matches 1. .312 of consensus"  
 repeat\_region 47915. .48083  
 /note="AluSq repeat: matches 145. .313 of consensus"  
 repeat\_region 49097. .49181  
 /note="L2 repeat: matches 2625. .2710 of consensus"  
 repeat\_region 49578. .49758  
 /note="MIR repeat: matches 9. .188 of consensus"  
 repeat\_region 49847. .49873  
 /note="MIR repeat: matches 155. .182 of consensus"  
 repeat\_region 49874. .50188  
 /note="AluB repeat: matches 3. .311 of consensus"  
 repeat\_region 50189. .50347  
 /note="MIR repeat: matches 5. .155 of consensus"  
 repeat\_region 51677. .51977  
 /note="AluSc repeat: matches 1. .303 of consensus"  
 repeat\_region 52025. .52123  
 /note="L2 repeat: matches 1655. .1753 of consensus"  
 repeat\_region 54337. .54544  
 /note="MIR repeat: matches 25. .246 of consensus"  
 repeat\_region 54754. .55032  
 /note="AluSp repeat: matches 1. .302 of consensus"  
 repeat\_region 55042. .55343  
 /note="AluY repeat: matches 5. .298 of consensus"  
 repeat\_region 56027. .56232  
 /note="MIR repeat: matches 21. .259 of consensus"  
 repeat\_region 56233. .56286  
 /note="18 copies 3 mer tgt 72% conserved"  
 repeat\_region 56342. .56673  
 /note="LIMA9 repeat: matches 5893. .6307 of consensus"  
 repeat\_region 56750. .56889  
 /note="L2 repeat: matches 2565. .2710 of consensus"  
 repeat\_region 57406. .57567  
 /note="MER5A repeat: matches 4. .187 of consensus"  
 repeat\_region 57864. .58174  
 /note="AluSq repeat: matches 1. .311 of consensus"  
 repeat\_region 59443. .60004  
 /note="L2 repeat: matches 2088. .2723 of consensus"  
 repeat\_region 60201. .60355  
 /note="MIR repeat: matches 20. .183 of consensus"  
 repeat\_region 63991. .64159  
 /note="MIR repeat: matches 78. .243 of consensus"  
 repeat\_region 64590. .64653  
 /note="32 copies 2 mer ta 70% conserved"  
 repeat\_region 64593. .64656  
 /note="16 copies 4 mer atat 68% conserved"  
 repeat\_region 64981. .65042  
 /note="MER5A repeat: matches 15. .72 of consensus"  
 repeat\_region 65043. .65341  
 /note="AluSc repeat: matches 1. .299 of consensus"  
 repeat\_region 65342. .65441  
 /note="MER5A repeat: matches 72. .186 of consensus"  
 repeat\_region 66825. .67131  
 /note="AluSc repeat: matches 1. .306 of consensus"

Query Match 44.5%; Score 159; DB 9; Length 129608;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-81;  
 Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 199 ttttgccagctgaggttctgtgtggaagacacctcactttcagaagaacaaca 258  
 |||||  
 Db 120383 TGTGGCCCTACGCTGAGGTGCTGCTGTGGAGAACCTCCTACTTTCAGAGAGACAACA 120324  
 |||||  
 Qy 259 gtaagcttggtttttcagcagcgggggtctctcatctttttttgtgtgtgagtt 318

```

|||||
Db 120323 GTAAGCTGGTGTTCACGACGGGGGTCTCTCATTTTCTTGGTTGAGTT 120264
QY 319 ggggattgagagggagggaggaagctgtgttg 357
|||||
Db 120263 GGGGATGGAGGAGGAGGAGGAAGGAAGTGTGTTG 120225

RESULT 6
AC012230/c
LOCUS
DEFINITION Homo sapiens clone RP11-IM10, WORKING DRAFT SEQUENCE, 39 unordered
pieces.
ACCESSION AC012230
VERSION AC012230.3 GI:7637254
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 175064)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens, clone RP11-IM10
Unpublished
2 (bases 1 to 175064)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collamore,A.,
Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Hearford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczy,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tefaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (21-OCT-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 22, 2000 this sequence version replaced gi:6454033.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L2510
Center clone name: 1.M.10
----- Summary Statistics
Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 117571 bases at least Q40
Consensus quality: 145749 bases at least Q30
Consensus quality: 160940 bases at least Q20
Insert size: 185000; agarose-fp
Insert size: 171264; sum-of-ctigs
Quality coverage: 2.9 in Q20 bases; agarose-fp
Quality coverage: 3.2 in Q20 bases; sum-of-ctigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 39 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will

```

\* be preserved.

1 1003: contig of 1003 bp in length

1004 1103: gap of 100 bp

1104 2634: contig of 1531 bp in length

2635 2734: gap of 100 bp

2735 4415: contig of 1681 bp in length

4416 4515: gap of 100 bp

4516 5785: contig of 1270 bp in length

5786 5885: gap of 100 bp

5886 7879: contig of 1994 bp in length

7880 7979: gap of 100 bp

7980 9686: contig of 1707 bp in length

9687 9786: gap of 100 bp

9787 12253: contig of 2467 bp in length

12254 12353: gap of 100 bp

12354 15228: contig of 2875 bp in length

15229 15328: gap of 100 bp

15329 17200: contig of 1872 bp in length

17201 17300: gap of 100 bp

17301 20131: contig of 2831 bp in length

20132 20231: gap of 100 bp

20232 22587: contig of 2356 bp in length

22588 22687: gap of 100 bp

22688 25707: contig of 3020 bp in length

25708 25807: gap of 100 bp

25808 28184: contig of 2377 bp in length

28185 28284: gap of 100 bp

28285 31338: contig of 3054 bp in length

31339 31438: gap of 100 bp

31439 34299: contig of 2861 bp in length

34300 34399: gap of 100 bp

34400 38318: contig of 3919 bp in length

38319 38418: gap of 100 bp

38419 42835: contig of 4417 bp in length

42836 42935: gap of 100 bp

42936 45448: contig of 2513 bp in length

45449 45548: gap of 100 bp

45549 48116: contig of 2568 bp in length

48117 48216: gap of 100 bp

48217 52618: contig of 4402 bp in length

52619 52718: gap of 100 bp

52719 56592: contig of 3874 bp in length

56593 56692: gap of 100 bp

56693 59635: contig of 2943 bp in length

59636 59735: gap of 100 bp

59736 63661: contig of 3926 bp in length

63662 63761: gap of 100 bp

63762 68437: contig of 4676 bp in length

68438 68537: gap of 100 bp

68538 71458: contig of 2921 bp in length

71459 71558: gap of 100 bp

71559 76888: contig of 5330 bp in length

76889 76988: gap of 100 bp

76989 82113: contig of 5125 bp in length

82114 82213: gap of 100 bp

82214 88220: contig of 6007 bp in length

88221 88320: gap of 100 bp

88321 93499: contig of 5179 bp in length

93500 93599: gap of 100 bp

93600 97901: contig of 4302 bp in length

97902 98001: gap of 100 bp

98002 103016: contig of 5015 bp in length

103017 103116: gap of 100 bp

103117 109178: contig of 6062 bp in length

109179 109278: gap of 100 bp

109279 117307: contig of 8029 bp in length

117308 117407: gap of 100 bp

117408 124079: contig of 6672 bp in length

124080 124179: gap of 100 bp

124180 131281: contig of 7102 bp in length

131282 131381: gap of 100 bp

131382 138059: contig of 6678 bp in length

138060 138159: gap of 100 bp

\* 138160 145491: contig of 7332 bp in length  
 \* 145492 145591: gap of 100 bp  
 \* 145592 157391: contig of 11800 bp in length  
 \* 157392 157491: gap of 100 bp  
 \* 157492 175064: contig of 17573 bp in length.

## FEATURES

source

Location/Qualifiers  
 1. .175064  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="RP11-1M10"  
 /clone\_lib="RPC1-11 Human Male BAC"  
 1. .1003  
 /note="assembly\_fragment"  
 1104. .2634  
 /note="assembly\_fragment"  
 2735. .4415  
 /note="assembly\_fragment"  
 4516. .5785  
 /note="assembly\_fragment"  
 5886. .7879  
 /note="assembly\_fragment"  
 7980. .9686  
 /note="assembly\_fragment"  
 9787. .12253  
 /note="assembly\_fragment"  
 12354. .15228  
 /note="assembly\_fragment"  
 15329. .17200  
 /note="assembly\_fragment"  
 17301. .20131  
 /note="assembly\_fragment"  
 20232. .22587  
 /note="assembly\_fragment"  
 22688. .25707  
 /note="assembly\_fragment"  
 25808. .28184  
 /note="assembly\_fragment"  
 28285. .31338  
 /note="assembly\_fragment"  
 31439. .34299  
 /note="assembly\_fragment"  
 34400. .38318  
 /note="assembly\_fragment"  
 38419. .42835  
 /note="assembly\_fragment"  
 42936. .45448  
 /note="assembly\_fragment"  
 45549. .48116  
 /note="assembly\_fragment"  
 48217. .52618  
 /note="assembly\_fragment"  
 52719. .56592  
 /note="assembly\_fragment"  
 56693. .59635  
 /note="assembly\_fragment"  
 59736. .63661  
 /note="assembly\_fragment"  
 63762. .68437  
 /note="assembly\_fragment"  
 68538. .71458  
 /note="assembly\_fragment"  
 clone\_end:sp6  
 vector\_side:left  
 71559. .76888

Query Match 44.5%; Score 159; DB 2; Length 175064;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-81;  
 Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 tgttgccctcagctgaggtgtgtgtggaagaacctcatttcagaagaacaaaca 258  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 151314 TGTGGCCTCAGCTGAGGTGCTGCTGTGGAAGAACCTCATTTCAGAAGACAAACA 151255

QY 259 gtaagcttggttttttcagcagcgggggtctctcatttttttttgggtttgagtt 318  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 151254 GTAAGCTTGGGTTTTCAGCAGCGGGGGTCTCTCAATTTTCTTGGTGGTTGAGTT 151195  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 319 ggggattggaggaggaggaggaggaggaggagctgtgttg 357  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 151194 GGGGATTGGAGGAGGAGGAGGAGGAGGAAGCTGTGTTG 151156  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

## RESULT

7

AX092589 183999 bp DNA linear PAT 21-MAR-2001  
 LOCUS  
 DEFINITION Sequence 1 from Patent WO0115676.  
 ACCESSION AX092589  
 VERSION AX092589.1 GI:13444647  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 183999)  
 Hayden.M.R.; Brooks-Wilson.A.R.; Pimstone.S.N. and Clee.S.M.  
 Compositions and methods for modulating hdl cholesterol and  
 triglyceride levels  
 Patent: WO 0115676-A 1 08-MAR-2001;  
 University of British Columbia (CA) ; Xenon Genetics Inc. (CA)  
 FEATURES  
 source  
 1. .183999  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 BASE COUNT 49549 a 37944 c 41170 g 54950 t 386 others  
 ORIGIN

Query Match 44.5%; Score 159; DB 6; Length 183999;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-81;  
 Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 tgttgccctcagctgaggtgtgtgtggaagaacctcatttcagaagaacaaaca 258  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 53328 TGTGGCCTCAGCTGAGGTGCTGCTGTGGAAGAACCTCATTTCAGAAGACAAACA 53387  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 259 gtaagcttggttttttcagcagcgggggtctctcatttttttttgggtttgagtt 318  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 53388 GTAAGCTTGGGTTTTCAGCAGCGGGGGTCTCTCAATTTTCTTGGTGGTTGAGTT 53447  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

QY 319 ggggattggaggaggaggaggaggaggagctgtgttg 357  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 53448 GGGGATTGGAGGAGGAGGAGGAGGAGGAAGCTGTGTTG 53486  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

## RESULT

8

AF287262 201144 bp DNA linear PRI 29-APR-2001  
 LOCUS  
 DEFINITION Homo sapiens ATP-binding cassette 1 sub-family A member 1 (ABCA1)  
 and SNAP protein genes, complete cds.  
 ACCESSION AF287262  
 VERSION AF287262.1 GI:13876612  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 201144)  
 Qiu,Y., Cavellier,L., Chiu,S., Yang,X., Rubin,E. and Cheng,J.F.  
 Human and mouse abcal comparative sequencing and transgenesis  
 studies revealing novel regulatory sequences  
 Genomics 73 (1), 66-76 (2001)  
 2 (bases 1 to 201144)  
 Qiu,Y., Cavellier,L., Chiu,S., Rubin,E. and Cheng,J.F.  
 Direct Submission  
 Submitted (13-JUL-2000) Genome Science Department, Lawrence

Query Match 44.5%; Score 159; DB 2; Length 175064;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-81;  
 Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 tgttgccctcagctgaggtgtgtgtggaagaacctcatttcagaagaacaaaca 258  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 151314 TGTGGCCTCAGCTGAGGTGCTGCTGTGGAAGAACCTCATTTCAGAAGACAAACA 151255

Berkeley National Laboratory, 1 Cyclotron Rd, MS 84-171, Berkeley, CA 94720, USA

FEATURES  
source

Location/Qualifiers  
1. .201144

/organism="Homo sapiens"  
/db\_xref="taxon:9606"

join(33931..34151,58319..58476,72893..72986,77520..77661,78931..79049,101048..101169,104152..104328,117287..117379,123229..123259,125285..125424,125754..125870,130079..130276,131023..131228,131749..131925,133708..133930,135677..135898,136959..137163,138268..138381,140179..140350,141340..141471,142777..142919,143123..143260,143963..144183,145441..145513,146657..146859,148527..148575,148771..148884,150280..150428,152078..152202,153438..153536,156568..156757,158278..158372,159682..159714,160837..160942,162417..162491,163013..163182,164413..164590,166588..166703,166815..166959,168505..168628,169715..169844,170109..170229,171017..171079,171987..172093,174448..174589,174961..175095,176039..176142,176625..176717,177376..177619,178560..181457)  
/gene="ABCA1"

/product="ATP-binding cassette 1 sub-family A member 1"

33931..181457

/gene="ABCA1"

33931..34151

/gene="ABCA1"

/number=1

58319..58476

/gene="ABCA1"

/number=2

join(58411..58476,72893..72986,77520..77661,78931..79049,101048..101169,104152..104328,117287..117379,123229..123259,125285..125424,125754..125870,130079..130276,131023..131228,131749..131925,133708..133930,135677..135898,136959..137163,138268..138381,140179..140350,141340..141471,142777..142919,143123..143260,143963..144183,145441..145513,146657..146859,148527..148575,148771..148884,150280..150428,152078..152202,153438..153536,156568..156757,158278..158372,159682..159714,160837..160942,162417..162491,163013..163182,164413..164590,166588..166703,166815..166959,168505..168628,169715..169844,170109..170229,171017..171079,171987..172093,174448..174589,174961..175095,176039..176142,176625..176717,177376..177619,178560..178700)  
/gene="ABCA1"

/codon\_start=1

/product="ATP-binding cassette 1 sub-family A member 1"

/protein\_id="AAK43526.1"

/db\_xref="GI:13876613"

/translation="MACPOLRLMLMKNLFRRTQCLLELVAVPLIFLILSVRL  
SYPPGEHQFNKAMPASGATLPWQGIICNANNPCFRPTGEAPGVGNPKSVIL  
ARLFSDDARLLLYSOKDTSKMDKVLRTLOQIKSSNLKLODFLVNETHFGFLYH  
NLSPKSTVDKMLRADVILHKVFLQYQLHLTSLCNGSKSEMIQDQDQVSLCLPL  
KEXLAARVLRNSMDLLKFLRLTNSPPSKELAEATKLLHSLGTLAQELFSMR  
SWSDRQVFLTNVSSSTQIYQAVSRVCGHPGGGLKISLNWEDNNYKALF  
GGNGTEDEATFYDNTTPYCNLMKLESSPLSIWIKALPLLVKILYTPDTPAT  
ROYMAEVNFTQELAVFLHDLGHEWELSPKWTFMENSOEMLVRLMLDSRONHFE  
QOLDGLDWAQIDVAFKHPEDVQSSNGSVYVREAFNETQAIITISRPMECVNLN  
KLPIATEVNLINKSMELLDERKFWAGIVFTGTPGSIPLPHVTKIRMDINDVERT  
NKIKDGVDPGPRADPFEDRMRYVVGFAVLDQVVEGAIIRVLTGTEKTKGVYMQMPY  
PCYVDDILFRNSRMLPMTLAWIYSAVIVKIVYEKARLKETMRIMGLDNLIL  
FSWFISSLPLLVISAGLLVILKGNLLPYSDPVVFLSVFAVVIILQCFLISLFL  
SRANLAACGGIITFYLYPLVLCVAVQYVGTFLKIFASLLSPVAFGCGCCFALFE  
EQIGVQWDLNFPESVEDGFLNLTSVSMLENTFLYGVMTWYIEAVFPQVQIRPPW  
YFPCTKSYNFGESDEKSHSGNKRISLTCHEEPETHLKLGVSONLVKIVYRDMKV  
AVDGLALNFYEGOITSLGHNAGKTTMSILTGFPPTSGTAYILGKDIRSEMSTIR  
ONLGVCFQHNVLFDMLTVEHFWYARLKLKSKHVAEQMALDGLPSSKLKSKT  
SOLSGMQRKSLAFVGGSKVILDEPTAGVDVPSRREGIILLKVRQGTIILST  
HMDADVLGDRIATIIHSGKLCVCSGLFLKQLNGTGYTLVKKDVSSLSRCSNNS

STVSVLKEDSVSSSDAGLSHSDTLTIDVSAISNLIRKIVSEARLVEDIGHEL  
TYVLPYEAAREGAFVELFHEIDRLSDIGSSYIGSETILEEIPFKVAESGVDAETS  
DGLTPARNRRRAFADKOSCLUPFTEDDAADPNDSIDIPESRETDLISGMDGKGYQVK  
GWLTKQOQFVALLKWRLLIARRSRKFFAIVLPAFVFCIALVSLIIVPPGKYSYKLE  
LQPMWYNEQYTFVNDAPEDTGTLELLNALT KDPGFCRMEGNPIPTDTCQAGEEM  
TTAPVPTIMDLFONGNWTQNPSPACQSSDKIKMLPVCPPGAGGLPPQKQNTA  
DILQDLGRNISDYLVKTYVQIIAKSLKNIWNERFYGGSLGVSNQALPPSQEYN  
DAIKMKHLKLIADSSADRFLNSLGRFTGLTRNNVYKWFNNKWHALSSFLNIN  
NAILRANLQGENSHYGTAFNHPNLTKQOLSEVALMTTSVDLVISICVIFAMSFV  
PASVVFILQIERSVSKARHLQFISGKPVIIYLSNFMVDMCNVVPATLVIIIFICQO  
KSYVSNLPLVALLLLLLYGMSITPLMPASFVKIPSTAYVVLTSVNLRTGINGSVA  
TFVLEFTDNKLNINDILAKSVFLIFPHFCGLRGLIDMKVNAKLSPLNDEDEDVRRRO  
RILDGGQNDILETKETKLYRKRKPVDRIKIPGECFGLLVNGAGKSKSTFKM  
LTGDTYTRGDAFLNKSILSNIEHVNMGYICQPDALTTELLTGREHVEFFALLRGV  
PEKEVGWEMARKKGLVYKERYAGNYSNGNKKLSTAMALIGGPPVVFLEDEPTTG  
MDPKARFLNCAVSVMYKGRSVVLTSHSMECEALCTRNAIMVNGFRCLGVSQHLK  
NREGDGYTIVRIAGSNPDLKVPQDFGLAFPGSVLKEKHNNLQVLPSSLSLARI  
FSILSOSKRLHIEDYSVSQTTLDQVFNFAKQSDDDHLKDLSLHKNQTVVDVAIT  
SELDQEKVKESYV"  
72893..72986  
/gene="ABCA1"  
/number=3  
77520..77661  
/gene="ABCA1"  
/number=4  
78931..79049  
/gene="ABCA1"  
/number=5  
101048..101169  
/gene="ABCA1"  
/number=6  
104152..104328  
/gene="ABCA1"  
/number=7  
117287..117379  
/gene="ABCA1"  
/number=8  
123229..122569  
/gene="ABCA1"  
/number=9  
125285..125424  
/gene="ABCA1"  
/number=10  
125754..125870  
/gene="ABCA1"  
/number=11  
130079..130276  
/gene="ABCA1"  
/number=12  
131023..131228  
/gene="ABCA1"  
/number=13  
131749..131925  
/gene="ABCA1"  
/number=14  
133708..133930  
/gene="ABCA1"  
/number=15  
135677..135898  
/gene="ABCA1"  
/number=16  
136959..137163  
/gene="ABCA1"  
/number=17  
138268..138381  
/gene="ABCA1"  
/number=18  
140179..140350  
/gene="ABCA1"  
/number=19  
141340..141471  
/gene="ABCA1"

exon

exon

exon

exon

exon

exon

exon

exon

exon

exon

exon

exon

exon

exon

exon

exon

exon

exon





16778	17678:	contig of 901 bp in length
17679	17778:	gap of 100 bp
17779	18679:	contig of 901 bp in length
18680	18779:	gap of 100 bp
18780	19632:	contig of 853 bp in length
19633	19732:	gap of 100 bp
19733	20634:	contig of 902 bp in length
20635	20734:	gap of 100 bp
20735	21620:	contig of 886 bp in length
21621	21720:	gap of 100 bp
21721	22579:	contig of 859 bp in length
22580	22679:	gap of 100 bp
22680	23558:	contig of 889 bp in length
23569	23668:	gap of 100 bp
23669	24554:	contig of 886 bp in length
24555	24654:	gap of 100 bp
24655	25521:	contig of 867 bp in length
25522	25621:	gap of 100 bp
25622	26487:	contig of 866 bp in length
26488	26587:	gap of 100 bp
26588	27464:	contig of 877 bp in length
27465	27564:	gap of 100 bp
27565	28466:	contig of 902 bp in length
28467	28566:	gap of 100 bp
28567	29464:	contig of 898 bp in length
29465	29564:	gap of 100 bp
29565	30447:	contig of 883 bp in length
30448	30547:	gap of 100 bp
30548	31453:	contig of 906 bp in length
31454	31553:	gap of 100 bp
31554	32452:	contig of 899 bp in length
32453	32552:	gap of 100 bp
32553	33447:	contig of 895 bp in length
33448	33547:	gap of 100 bp
33548	34435:	contig of 888 bp in length
34436	34535:	gap of 100 bp
34536	35433:	contig of 898 bp in length
35434	35533:	gap of 100 bp
35534	36440:	contig of 907 bp in length
36441	36540:	gap of 100 bp
36541	37422:	contig of 882 bp in length
37423	37522:	gap of 100 bp
37523	38402:	contig of 880 bp in length
38403	38502:	gap of 100 bp
38503	39380:	contig of 878 bp in length
39381	39480:	gap of 100 bp
39481	40407:	contig of 927 bp in length
40408	40507:	gap of 100 bp
40508	41405:	contig of 898 bp in length
41406	41505:	gap of 100 bp
41506	42399:	contig of 894 bp in length
42400	42499:	gap of 100 bp
42500	43357:	contig of 858 bp in length
43358	43457:	gap of 100 bp
43458	44356:	contig of 899 bp in length
44357	44456:	gap of 100 bp
44457	45325:	contig of 869 bp in length
45326	45425:	gap of 100 bp
45426	46305:	contig of 880 bp in length
46306	46405:	gap of 100 bp
46406	47302:	contig of 897 bp in length
47303	47402:	gap of 100 bp
47403	48293:	contig of 891 bp in length
48294	48393:	gap of 100 bp
48394	49280:	contig of 887 bp in length
49281	49380:	gap of 100 bp
49381	50250:	contig of 870 bp in length
50251	50350:	gap of 100 bp
50351	51209:	contig of 859 bp in length
51210	51309:	gap of 100 bp
51310	52221:	contig of 912 bp in length
52222	52321:	gap of 100 bp
52322	53205:	contig of 884 bp in length

*	53206	53305:	gap of	100 bp
*	53306	54225:	contig of 920 bp	in length
*	54226	54325:	gap of	100 bp
*	54326	55206:	contig of 881 bp	in length
*	55207	55306:	gap of	100 bp
*	55307	56165:	contig of 859 bp	in length
*	56166	56295:	gap of	100 bp
*	56266	57124:	contig of 859 bp	in length
*	57125	57224:	gap of	100 bp
*	57225	58121:	contig of 897 bp	in length
*	58122	58221:	gap of	100 bp
*	58222	59075:	contig of 854 bp	in length
*	59076	59175:	gap of	100 bp
*	59176	60058:	contig of 883 bp	in length
*	60059	60138:	gap of	100 bp
*	60159	61067:	contig of 909 bp	in length
*	61068	61167:	gap of	100 bp
*	61168	62051:	contig of 884 bp	in length
*	62052	62151:	gap of	100 bp
*	62152	63022:	contig of 871 bp	in length
*	63023	63122:	gap of	100 bp
*	63123	84046:	contig of 904 bp	in length
*	64027	64126:	gap of	100 bp
*	64127	65015:	contig of 889 bp	in length
*	65016	65115:	gap of	100 bp
*	65116	65975:	contig of 860 bp	in length
*	65976	66075:	gap of	100 bp
*	66076	66954:	contig of 879 bp	in length
*	66955	67054:	gap of	100 bp
*	67055	67947:	contig of 893 bp	in length
*	67948	68047:	gap of	100 bp
*	68048	68934:	contig of 887 bp	in length
*	68935	69034:	gap of	100 bp
*	69035	69910:	contig of 876 bp	in length
*	69911	70010:	gap of	100 bp

Query Match 32.5%; Score 116; DB 2; Length 90698;  
Best Local Similarity 100.0%; Pred. No. 3.7e-56;  
Matches 116; Conservative 0; Mismatches 0; Indels 0

Qy	1	tggaggctcagctgagaggctggtatgacgctcctcattggtgtaggtttgcagca	60
Db	13169	TGGAGGCTCAGCTGAGAGGGCTGGATTAGCAGTCCCTATTGGTGTATGGCTTTCAGCA	13110
Qy	61	ataactgatgctgtttccctcctgcttatcttcagttaatgacagcagcg	116
Db	13109	ATAACTGATGCTGTTTCCCTCCTCGTGTATCTTTTCAGTTAATGACAGCAGCGG	13054

RESULT 10

AX351035	AX351035	99 bp	DNA	linear	PAT 06-FEB-2002
LOCUS	Sequence 7 from Patent WO0183746.				
DEFINITION	AX351035				
ACCESSION	AX351035				
VERSION	AX351035.1	GI:18616391			
KEYWORDS	.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (sites)				
AUTHORS	Rosler-Montus,M.F., Prades,C., Lemoine,C., Naudin,L., Denefle,P., Breuer,B., Duverger,N., Remaley,A. and Santamarina-Fojo,S.				
TITLE	Regulatory nucleic acid sequences of the abcl gene				
JOURNAL	Patent: WO 0183746-A 7 08-NOV-2001;				
	Avantis Pharma S.A. (FR)				

BASE COUNT	ORIGIN
17 a	22 c
17 g	27 g
33 t	

Query Match 27.7%; Score 99; DB 6; Length 99;  
Best Local Similarity 100.0%; Pred. No. 2.4e-46;  
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tggaggtctcagctgagagggctggtgattagcagctcattggtgtatgcttttcagca 60  
Db 1 TGGAGGTCTCAGCTGAGAGGGCTGGATTAGCAGCTCCTGCTGTATGCTGTTCAGCA 60

QY 61 ataactgagctgtttccctcctcttattcttcag 99  
Db 61 ATAACATGATGGCTGTTTCCCTCCTGCTTAACTTTCAG 99

RESULT 11  
AX351036 AX351036 99 bp DNA linear PAT 06-FEB-2002  
LOCUS Sequence 8 from Patent WO0183746.  
DEFINITION AX351036  
ACCESSION AX351036  
VERSION AX351036.1 GI:18616392  
KEYWORDS human.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (sites)  
AUTHORS Rosier-Montus, M.F., Prades, C., Lemoine, C., Naudin, L., Denefle, P.,  
Brewer, B., Duverger, N., Remaley, A., and Santamarina-Fojo, S.  
TITLE Regulatory nucleic acid sequences of the abcl gene  
JOURNAL Patent: WO 0183746-A 8 NOV-2001;  
Aventis Pharma S.A. (FR)  
FEATURES Location/Qualifiers  
source 1..99  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 15 a 9 c 41 g 34 t  
ORIGIN

Query Match 27.7%; Score 99; DB 6; Length 99;  
Best Local Similarity 100.0%; Pred. No. 2.4e-46;  
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 259 gtagcttggttttcagcagcgggggtctctcatttttttctgtggttttgagtt 318  
Db 1 GTAAGCTTGGTGTTCAGCAGCGGGGGTCTCTCATTTTCTTGTGTTGAGTT 60

QY 319 gggattgagagggaggagggaaggaagctgtgtg 357  
Db 61 GGGATTGGAGGAGGAGGAGGAGGAGGAGGAGCTGTGTG 99

RESULT 12  
AF258623S2 AF258623S2 200 bp DNA linear PRI 23-JUN-2000  
LOCUS Homo sapiens ATP binding cassette transporter 1 (ABCA1) gene, exon  
DEFINITION 2.  
ACCESSION AF258624  
VERSION AF258624.1 GI:7769714  
KEYWORDS 2 of 4  
SEGMENT human.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 200)  
AUTHORS Pullinger, C.R., Hakamata, H., Duchateau, P.N., Eng, C.,  
Auizeraat, B.E., Fielding, C.J., and Kane, J.P.  
TITLE Analysis of hABC1 gene 5' end: additional peptide sequence,  
promoter region, and four polymorphisms  
JOURNAL Biochem. Biophys. Res. Commun. 271 (2000) In press  
REFERENCE 2 (bases 1 to 200)  
AUTHORS Pullinger, C.R., Hakamata, H., Duchateau, P.N., Eng, C.,

Auizeraat, B.E., Fielding, C.J. and Kane, J.P.  
Direct Submission  
Submitted (23-JUN-2000) Cardiovascular Research Institute,  
University of California, San Francisco, 505 Parnassus Avenue, San  
Francisco, CA 94143-0130, USA  
Sequence update by submitter  
REMARK Location/Qualifiers  
FEATURES source 1..200  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="9"  
/map="9q31"  
22..179  
/gene="ABCA1"  
/number=2  
BASE COUNT 37 a 56 c 57 g 50 t  
ORIGIN

Query Match 22.7%; Score 81; DB 9; Length 200;  
Best Local Similarity 100.0%; Pred. No. 8.5e-36;  
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 tttggcttcagctgaggtgtgtgtggaagacccctcactttcagaagaagacaaca 258  
Db 120 TGTGGCTCAGCTGAGGTGTGCTGTGCTGTGGAAGAACCTCCTTTCAGAGAAGACAAACA 179

QY 259 gtaagcttggttttttcagca 279  
Db 180 GTAAGCTTGGGTTTTTCAGCA 200

RESULT 13  
AX060713 AX060713 10442 bp DNA linear PAT 22-JAN-2001  
LOCUS Sequence 1 from Patent WO0078972.  
DEFINITION AX060713  
ACCESSION AX060713  
VERSION AX060713.1 GI:12406103  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 10442)  
AUTHORS Lawn, R.M., Wade, D. and Garvin, M.  
TITLE Regulation with binding cassette transporter protein abcl  
JOURNAL Patent: WO 0078972-A 1 28-DEC-2000;  
CV THERAPEUTICS, INC. (US)  
FEATURES Location/Qualifiers  
source 1..10442  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 2898 a 2297 c 2408 g 2835 t 4 others  
ORIGIN

Query Match 21.6%; Score 77; DB 6; Length 10442;  
Best Local Similarity 99.2%; Pred. No. 2.2e-33;  
Matches 127; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 131 agcttgccgcgtgccttcagggctcccgagccacacgcgtggcgctgctgagggga 190  
Db 229 AGCTCTGGCGCTGCCCTTCCAGGGCTCCCGAGGCACACGCTGGGCGTGTGCTGAGGA 288

QY 191 acatggcatgtggcctcagctgaggtgtgtgtgtggaagacccctcactttcagaaga 250  
Db 289 ACATGGCTTGTGGCCTCAGCTGAGGTGTGTGTGTGGAAGACCTCCTTTCAGAGAA 348

QY 251 gacaaaca 258  
Db 349 GACAAACA 356

```

RESULT 14
AX060892
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
Source
BASE COUNT
ORIGIN
Query Match
Best Local Similarity
Matches 127; Conservative
QY 131 agctctggcgcgtgctctccagggctccgagcagcacgctggcgctgctgagggga 190
Db 229 AGCTCTGGCGCTGCTTCCAGGGCTCCGAGCAGCAGCTGGCGTGGCTGAGGGA 288
QY 191 acatggcatgttgctcagctcagctgaggtgctgctgctggaagaacctcacttcagaaga 250
Db 289 ACATGGCTTGTGGCTCAGCTGAGGTGCTGCTGCTGGAAGAAGCTCCTTTCAGAGAA 348
QY 251 gacaaaca 258
Db 349 GACAAACA 356
RESULT 15
AF285167
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
Source
BASE COUNT
ORIGIN
Query Match
Best Local Similarity
Matches 127; Conservative
QY 131 agctctggcgcgtgctctccagggctccgagcagcacgctggcgctgctgagggga 190
Db 229 AGCTCTGGCGCTGCTTCCAGGGCTCCGAGCAGCAGCTGGCGTGGCTGAGGGA 288
QY 191 acatggcatgttgctcagctcagctgaggtgctgctgctggaagaacctcacttcagaaga 250
Db 289 ACATGGCTTGTGGCTCAGCTGAGGTGCTGCTGCTGGAAGAAGCTCCTTTCAGAGAA 348
QY 251 gacaaaca 258
Db 349 GACAAACA 356

```

CDS

```

/gene="ABCA1"
291..7076
/gene="ABCA1"
/codon_start=1
/product="ATP-binding cassette transporter 1"
/protein_id="AAP98175.1"
/db_xref="GI:9755159"
/translation="MACWPQLRLLLMKNTFRROTCCQLLELVAMPFLFLLILSVRL
SYYPQHECHFFNKASAGTLPWQVGIICANNPCFRYPTGEAGVGVNFKSIV
ARLFSARLLLYSQRDTSKMDKRLVQLHQLKSSNKLQDFLDVNETSGGLYH
NLSPKSTVDKMLRADVLHVKVQGLHQLKSSNKLQDFLDVNETSGGLYH
KEKLAARVLRNMDILKPLILNSTSPSKELAEATKLLHSLGTLAQLFMSR
SWSDMRQEVMLTNVSSSTSIQASRIVCVCHGEGGLKSLNWTEDNNYKALF
GGNGTEDEATFYDNSTTTPYCNLMKNLESSPLRIWKLKPLLVKGLIYDPDPAT
ROYMAEVNKTQELAVHLEGGMEELSPKIVTMEASQEMDLVRLMSLRNDHFEW
QQLDGLWTAQDIAVFAKHPEDVQSSNGSVTREFAFNETQAIIRSFNECVNHLN
KLEPIATEVWLNKSMELDERKFAIVFTGTPGSIIEHPHVKYKIRMDINDVPT
NKIKDGYWDPGRADPFEDMRVWVGFAVLDVVEQAIIRVLTGTEKTKGVYQMPY
PCVVDIFLRVMSRSMPLFWLAWIYSVAIIIGIVYEKARLKTMRIMGLDNLILW
FSWFISLLIPLVLSAGLLVILKGNLLPYSGFVVFVFLSVFVAVVILQCFLISLF
SRANLAACGGIIIFTLYLPYVLCVAVWDYVGTFLKIFASLSPVAFGCGEYFALFE
EGGIGVQNDLFESEVEDGFNLTTSSIMLEDFTLGYVMTWYIEAFPGQYIGPRPW
YFPTKSYWFEESDEKSHPSNQRKMEICEEETHLKLGVSIQNLVKKVYDGMKV
AVDGLANFEGQITFSLHNGAGKTTMSILTLGLPPTSGTAYILGDIRSEMTIR
QNLGVCPOHNVLFDMLTVEHIFWYARLKLSEKHVKAEMQALDGLPSSKLKST
SOLSGMOKLSVALAFVGSKVILDEPTAGVDPSYSRGIWELLKLLKQGRITILST
HMDADVLGDRALIIHSGKLCVASSLFLKQLGTGYTLVKKDVESSLSCNRS
SVSYLKKEDSVSSQSDAGLDHSDTLTIDVSAISNLIRKHSEARLVEDIGHEL
TLYPYEAKAGAFVELFHEIDRLSDGISYSETLEIFLLKVAPESDVDAETS
DGTLPARRNRAGFGKQSLRPTEDDAADPDNDSDIPESRETLKSGDGSQVVK
GWLKIQOQVALLKRLIARRSRKFFAIVLPAVFCVIALVSLIYVPPFKYPSLE
LQPMVNEQYTFVSNDAEDTGTLELNLATKDPGCTRCMEGNIPDTPCOAGSEW
TTAPVPTIMDLFONGMTQNPSPACQSSDKIKMLPVCPPGAGGLPPQKQONTA
DILODGTENISDYLKTVQIIAASLKNHWNFEFRYGGFSLGVSNTQALPPSOEVN
DAIKOMKHLKLAKSSADRLNSLGRMTGLDRNNKWNFNNGHAISSFLAVIN
NAILRANOKGENSHYGFATFNHPLNTKQSLSEVALMTSDVLSVICFAMSFY
PASVYFLIQERVSKAHLQFISGVKPVYIWLSEFVMDMNTVVPATLVIIFCQQ
KSYVSTNLPVIALLLLYGWSITPLMPASVFKIPSTAYVLTSVNLTGINGSVA
TFVLELTDKNLNNIDILKSVFLIPHPCLGRGLIDMKVQAMADALERGENRVS
PLSWDLVGRNLFAMAVEGVVFLITVLOVREFIIPRPNNAKLSPLNDEDEVRERQ
RILDGGQNDLILEIKELPKIYRKRKPAVDICVGPGEFCGLLGVNGAGKSTFKM
LIGDPTVTRGDALFNKNSILSNIEHVHQNMGQPOFDAITELLTGREHVEFFALLRGV
PEKEKRGWEWAIRKGLVYKGYAGNYSGNKRKLSTAMALIGGPVPLDEPTTG
MDPKARRLWNCALSVKREGSVLTSMSMECEALCTRMALIMVNGRRCVGLSVOLK
NRFGGYTVVIRAGSNPLKPVODFFGLAFPGSVLKEKHNMLQYQLPSSLSIARI
FSILSQSKRLHIEDYSVSQITLDQVVFVNEAKDQSDDDHLKDLSLHKQTVVDVAULT
SFLODEKVKESVY"
BASE COUNT 2898 a 2297 c 2408 g 2835 t 4 others
ORIGIN

```

```

Query Match 21.6%; Score 77; DB 9; Length 10442;
Best Local Similarity 99.2%; Pred. No. 2.2e-33;
Matches 127; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 131 agctctggcgcgtgctctccagggctccgagcagcacgctggcgctgctgagggga 190
Db 229 AGCTCTGGCGCTGCTTCCAGGGCTCCGAGCAGCAGCTGGCGTGGCTGAGGGA 288
QY 191 acatggcatgttgctcagctcagctgaggtgctgctgctggaagaacctcacttcagaaga 250
Db 289 ACATGGCTTGTGGCTCAGCTGAGGTGCTGCTGCTGGAAGAAGCTCCTTTCAGAGAA 348
QY 251 gacaaaca 258
Db 349 GACAAACA 356
RESULT 16
AX060719
LOCUS
DEFINITION
ACCESSION
AX060719
Sequence 7 from Patent WO0078972.
linear
PAT 22-JAN-2001

```

VERSION AX060719.1 GI:12406108  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
TITLE 1 (bases 1 to 10474)  
JOURNAL Lawn,R.M., Wade,D., and Garvin,M.  
CV Regulation with binding cassette transporter protein abcl  
PATENT: WO 0078972-A 7 28-DEC-2000;  
CV THERAPEUTICS, INC. (US)  
FEATURES  
Location/Qualifiers  
1..10474  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 2906 a 2305 c 2416 g 2843 t 4 others  
ORIGIN  
Query Match 21.6%; Score 77; DB 6; Length 10474;  
Best Local Similarity 99.2%; Pred. No. 2.2e-33;  
Matches 127; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 131 agctctggccgtgctccagggctcccgagccacacgctggcgctgctgagggga 190  
|||||  
Db 261 AGCTCTGGCCGCTGCTCCAGGGCTCCCGAGCCACACGCTGGGGTGTGCTGAGGGA 320  
QY 191 acatggcatgttgccctcagctgaggttgctgctgtggaagaacctcacttcagaagaa 250  
|||||  
Db 321 ACATGGCTGTGTGGCTCAGCTGAGGTGCTGTGGAAGAACCCTCACTTTTCAAGAAGAA 380  
QY 251 gacaaaca 258  
|||||  
Db 381 GACAAACA 388  
RESULT 17  
AX060721  
LOCUS 10474 bp DNA linear PAT 22-JAN-2001  
DEFINITION Sequence 9 from Patent W00078972.  
ACCESSION AX060721  
VERSION AX060721.1 GI:12406109  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
TITLE 1 (bases 1 to 10474)  
JOURNAL Lawn,R.M., Wade,D., and Garvin,M.  
CV Regulation with binding cassette transporter protein abcl  
PATENT: WO 0078972-A 9 28-DEC-2000;  
CV THERAPEUTICS, INC. (US)  
FEATURES  
Location/Qualifiers  
1..10474  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 2907 a 2304 c 2415 g 2844 t 4 others  
ORIGIN  
Query Match 21.6%; Score 77; DB 6; Length 10474;  
Best Local Similarity 99.2%; Pred. No. 2.2e-33;  
Matches 127; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 131 agctctggccgtgctccagggctcccgagccacacgctggcgctgctgagggga 190  
|||||  
Db 261 AGCTCTGGCCGCTGCTCCAGGGCTCCCGAGCCACACGCTGGGGTGTGCTGAGGGA 320  
QY 191 acatggcatgttgccctcagctgaggttgctgctgtggaagaacctcacttcagaagaa 250  
|||||  
Db 321 ACATGGCTGTGTGGCTCAGCTGAGGTGCTGTGGAAGAACCCTCACTTTTCAAGAAGAA 380  
QY 251 gacaaaca 258  
|||||  
Db 381 GACAAACA 388

|||||  
Db 381 GACAAACA 388  
RESULT 18  
AX060898  
LOCUS 10474 bp DNA linear PAT 22-JAN-2001  
DEFINITION Sequence 7 from Patent W00078971.  
ACCESSION AX060898  
VERSION AX060898.1 GI:12406275  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
TITLE 1 (bases 1 to 10474)  
JOURNAL Lawn,R.M., Wade,D., Oram,J.F. and Garvin,M.  
CV Atp binding cassette transporter protein abcl polypeptides  
PATENT: WO 0078971-A 7 28-DEC-2000;  
CV THERAPEUTICS, INC. (US)  
FEATURES  
Location/Qualifiers  
1..10474  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 2906 a 2305 c 2416 g 2843 t 4 others  
ORIGIN  
Query Match 21.6%; Score 77; DB 6; Length 10474;  
Best Local Similarity 99.2%; Pred. No. 2.2e-33;  
Matches 127; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 131 agctctggccgtgctccagggctcccgagccacacgctggcgctgctgagggga 190  
|||||  
Db 261 AGCTCTGGCCGCTGCTCCAGGGCTCCCGAGCCACACGCTGGGGTGTGCTGAGGGA 320  
QY 191 acatggcatgttgccctcagctgaggttgctgctgtggaagaacctcacttcagaagaa 250  
|||||  
Db 321 ACATGGCTGTGTGGCTCAGCTGAGGTGCTGTGGAAGAACCCTCACTTTTCAAGAAGAA 380  
QY 251 gacaaaca 258  
|||||  
Db 381 GACAAACA 388  
RESULT 19  
AX060900  
LOCUS 10474 bp DNA linear PAT 22-JAN-2001  
DEFINITION Sequence 9 from Patent W00078971.  
ACCESSION AX060900  
VERSION AX060900.1 GI:12406276  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
TITLE 1 (bases 1 to 10474)  
JOURNAL Lawn,R.M., Wade,D., Oram,J.F. and Garvin,M.  
CV Atp binding cassette transporter protein abcl polypeptides  
PATENT: WO 0078971-A 9 28-DEC-2000;  
CV THERAPEUTICS, INC. (US)  
FEATURES  
Location/Qualifiers  
1..10474  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 2907 a 2304 c 2415 g 2844 t 4 others  
ORIGIN  
Query Match 21.6%; Score 77; DB 6; Length 10474;  
Best Local Similarity 99.2%; Pred. No. 2.2e-33;  
Matches 127; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 131 agctctggcgcgtcctccaggggtcccgagccacagctggcgctgctgagagga 190
Db 261 AGCTCTGGCGCGTCTGCTTCCAGGGTCCCGAGCCACAGCTGGCGTCTGGCTGAGGGA 320

QY 191 acatgcatgttgccctcagctcaggttgctgctgtggaagaacctcacttccagaagaa 250
Db 321 ACATGCTGTGTGGCGCTCAGCTGAGGTGCTGCTGTGGAAGAACCCTCACTTTCAAGAA 380

QY 251 gacaaaca 258
Db 381 GACAAACA 388

RESULT 20
AB037924
LOCUS Homo sapiens mRNA for ABC1, partial cds. PRI 12-OCT-2000
DEFINITION Homo sapiens mRNA for ABC1, partial cds.
ACCESSION AB037924
VERSION AB037924.1 GI:9711458
KEYWORDS ABC1.
SOURCE Homo sapiens placenta cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Zhao,L.X., Zhou,C.J., Tanaka,A., Nakata,M., Hirabayashi,T.,
Amachi,T., Shioda,S., Ueda,K. and Inagaki,N.
TITLE Cloning, characterization and tissue distribution of the rat
ATP-binding cassette (ABC) transporter ABC2/ABCA2
JOURNAL Biochem. J. 350 (Pt 3), 865-872 (2000)
PUBMED 10970803
REFERENCE 2 (bases 1 to 298)
AUTHORS Ueda,K., Kikuchi,N. and Tanaka,A.
TITLE Direct Submission
JOURNAL Submitted (02-FEB-2000) Kazumitsu Ueda, Kyoto University Graduate
School of Agriculture, Division of Applied Life Sciences;
Kitashirakawa, Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
(E-mail:uedak@kais.kyoto-u.ac.jp, Tel:81-75-753-6105,
Fax:81-75-753-6104)
FEATURES
source
1..298
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="placenta"
88..298
/gene="ABC1"
88..>298
/codon_start=1
/product="ABC1"
/protein_id="BAB07875.1"
/db_xref="GI:9711459"
/translation="MACWPQLRLLLKNTFRRRTQCQLLEVAWPLFIFLILISVRL
SYPPYQHECFHPKAMPAGTLPVY"
BASE COUNT 60 a 87 c 77 g 72 t 2 others
ORIGIN

Query Match 16.8%; Score 60; DB 9; Length 298;
Best Local Similarity 100.0%; Pred. No. 1.7e-23;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 tgttgccctcagctcaggttgctgctgtggaagaacctcacttccagaagaacaaaca 258
Db 94 TGTGGCGCTCAGCTGAGGTGCTGCTGTGGAAGAACCCTCACTTTCAAGAAACA 153

RESULT 21
AX127764
LOCUS AX127764 446 bp DNA linear PRI 15-MAY-2001
DEFINITION Sequence 3 from Patent WO0130848.
ACCESSION AX127764
VERSION AX127764.1 GI:14134411

```

```

KEYWORDS synthetic construct.
SOURCE synthetic construct
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 446)
AUTHORS Deneffe,P., Rosier-Montus,M.F., Arnould-Reguigne,I., Prades,C.,
Naudin,L., Lemoine,C., Duverger,N., Jaye,M., Searfoss,G.H.,
Remaley,A., Brewer,H.B. and Dean,M.
TITLE Nucleic acids of the human abcl gene and their therapeutic and
diagnostic application
JOURNAL Patent: WO 0130848-A 3 03-MAY-2001;
Aventis Pharma S.A. (FR)
FEATURES
source
1..446
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Oligonucleotide Primer"
BASE COUNT 96 a 123 c 112 g 115 t
ORIGIN

Query Match 16.8%; Score 60; DB 6; Length 446;
Best Local Similarity 100.0%; Pred. No. 1.7e-23;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 tgttgccctcagctcaggttgctgctgtggaagaacctcacttccagaagaacaaaca 258
Db 191 TGTGGCGCTCAGCTGAGGTGCTGCTGTGGAAGAACCCTCACTTTCAAGAAACA 250

RESULT 22
AX139751
LOCUS AX139751 446 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 3 from Patent EP1096012.
ACCESSION AX139751
VERSION AX139751.1 GI:14275333
KEYWORDS synthetic construct.
SOURCE synthetic construct
ORGANISM artificial sequence.
REFERENCE 1 (bases 1 to 446)
AUTHORS Deneffe,P., Rosier-Montus,M.F., Arnould-Reguigne,I., Prades,C.,
Naudin,L., Lemoine,C., Duverger,N., Jaye,M., Searfoss III,G.H.,
Remaley,A., Brewer,H.B. and Dean,M.
TITLE Nucleic acids of the human abcl gene and their therapeutic and
diagnostic application
JOURNAL Patent: EP 1096012-A 3 02-MAY-2001;
Aventis Pharma S.A. (FR)
FEATURES
source
1..446
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Oligonucleotide Primer"
BASE COUNT 96 a 123 c 112 g 115 t
ORIGIN

Query Match 16.8%; Score 60; DB 6; Length 446;
Best Local Similarity 100.0%; Pred. No. 1.7e-23;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 tgttgccctcagctcaggttgctgctgtggaagaacctcacttccagaagaacaaaca 258
Db 191 TGTGGCGCTCAGCTGAGGTGCTGCTGTGGAAGAACCCTCACTTTCAAGAAACA 250

RESULT 23
AF258627
LOCUS AF258627 697 bp mRNA linear PRI 11-MAY-2000
DEFINITION Homo sapiens ATP binding cassette transporter 1 (ABCA1) mRNA,
partial cds.
ACCESSION AF258627
VERSION AF258627.1 GI:7769707

```

KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS 1 (bases 1 to 697)  
Pullinger, C.R., Hakamata, H., Duchateau, P.N., Eng, C., Aouizerat, B.E., Fielding, C.J. and Kane, J.P.  
TITLE Analysis of hABC1 gene 5' end: additional peptide sequence, promoter region, and four polymorphisms  
JOURNAL Biochem. Biophys. Res. Commun. 271 (2000) In press  
REFERENCE 2 (bases 1 to 697)  
Pullinger, C.R., Hakamata, H., Duchateau, P.N., Eng, C., Aouizerat, B.E., Fielding, C.J. and Kane, J.P.  
TITLE Direct Submission  
JOURNAL Submitted (19-APR-2000) Cardiovascular Research Institute, University of California, San Francisco, 505 Parnassus Avenue, San Francisco, CA 94143-0130, USA  
FEATURES  
source Location/Qualifiers  
1..697  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="9"  
/map="9c31"  
/tissue\_type="placenta"  
1..>697  
/gene="ABCAL"  
/note="ABC1"  
396..>697  
/gene="ABCAL"  
/note="membrane-bound"  
/codon\_start=1  
/product="ATP binding cassette transporter 1"  
/protein\_id="AAF69513.1"  
/db\_xref="GI:7769708"  
/translation="MACWPOLRLLLKLNLFRRROTCOLLLEAVPFLIFILISVRL  
SYYPEQHECHFPNKAMPSAGTLPWQGIICNANPCFRYPPTGAPGVVGNFKS  
152 a 198 c 190 g 156 t 1 others  
BASE COUNT  
ORIGIN  
Query Match 16.8%; Score 60; DB 9; Length 697;  
Best Local Similarity 100.0%; Pred. No. 1.7e-23;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 199 ttttgccctcagctgaggttgcgtctgtggaagaccctcactttcagaagaagacaaaca 258  
|||||  
Db 402 TGTGGCCTCAGCTGAGGTGCTGCTGTGGAAGAACCTCACTTTTCAGAAGAAGACAAACA 461  
RESULT 24  
AB055982  
LOCUS Homo sapiens mRNA for ABCAL, complete cds. PRI 18-AUG-2001  
DEFINITION  
ACCESSION AB055982  
VERSION  
KEYWORDS  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS 1 (bases 1 to 6786)  
Tanaka, A.R., Abe-Dohmae, S., Arakawa, R., Sadanami, K., Kidera, A., Kioka, N., Amachi, T., Yokoyama, S. and Ueda, K.  
TITLE A new topological model of functional human ABCAL-Signal peptide cleavage and glycosylation of a large extracellular domain  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 6786)  
Ueda, K., Kioka, N. and Tanaka, A.R.  
AUTHORS Direct Submission  
TITLE Submitted (20-FEB-2001) Kazumitsu Ueda, Kyoto University Graduate School, Applied Life Sciences, Kitashirakawa, Kyoto Sakyo-ku, Kyoto 606-8502, Japan (E-mail:uedak@kais.kyoto-u.ac.jp).

Tel:81-75-753-6105, Fax:81-75-753-6104)  
FEATURES  
source Location/Qualifiers  
1..6786  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
1..6786  
/gene="ABCAL"  
1..6786  
/gene="ABCAL"  
/codon\_start=1  
/product="ABCAL"  
/protein\_id="BAB63210.1"  
/db\_xref="GI:15212107"  
/translation="MACWPOLRLLLKLNLFRRROTCOLLLEAVPFLIFILISVRL  
SYYPEQHECHFPNKAMPSAGTLPWQGIICNANPCFRYPPTGAPGVVGNFKS  
ARLFSADARLLYSOKDISGKMDKRVLTLOQIKSSNKLQDLVDNENFSGSLYH  
NLSLPKSTVDKMRADVLHKVLOGYLHSLCNGSEMIQGLQGVSELCGLP  
RKLAAARVLRNMDILKPIRLTNSFPFSEKLAETKLLHSLGLTQAEVSMR  
SWSMDRQEVFLTNVSSSSQIYQAVSRIVCGHPEGGLKILSNWYEDNNYKALF  
GNGTEEDAETFDNSTTTCNDLMLKNLESSPLRIWKALPLVLGKTLTPDTPAT  
QVMAEVNKTFOELAVFHDLGEMWELSPKIWTENSOEMDLVRLMLDSRDNHWE  
QDLGLDWTQADIVAFIAKHPEDVQSSNGSVYTWREAFNETQOALRTSRWECVNLN  
KLEIATEVNLINKSMELDERKFWAGIVFTGITPGSEIPLHHVYKIKRMDIDNVERT  
NKIKDGYWDPGRADPFDMDYVWGGFVAYLQDVEQAILRVLTGTEKGTVMYQMPY  
PCYVDDIFLRVMSRMPLEMTLAWIYSVAIIKIGVIEKEARLKETRMJLNDLSILM  
FSWISSLIPLVSLAGLVVILKLNLPYSDPSVVFVLSVFAVVTILQCELIISLF  
SRANLAAACGGIIFYTLVPLVLCVADQYVGTFLKIFASLLSPVAFPGQYIGRPW  
EQGIGVQNDNLFESPVEDGFLNLTSSVMMLEDTLYGYMTWYIEAVFPQYIGRPW  
YFPTKSYWFEESDEKSHPGNKRKISEICWEEETHLKLGVSQNLVYKVRDGMKV  
AVDGLANFYEGQITSLFNGHAGKTTMSILTLGLPPTSGTAYILGKDISEMSTIR  
QNLGVCQHNVLDFMLTBEHFWFVYARLGLSEKHVKAEMEQLDVLGPPSKLSKT  
SOLSGQMQRKLSVALAVGGSKVYILDEPTAGVDVPSRRGIWELLIKYRQRTIILST  
HMDADVLGDRATISHGKLCVGSLLFLKNQGLGYTLVKKDSESLSSCRNSS  
STVSYLKEDSVSSQSSDAGLGSDESFTLTIDVSAISNLIRKHVSEARLVEDIGHEL  
TYVLPYAAKEGAFELFHEIDRLSDLGISGSETTLEELFLKVAESGVDAETS  
DGTLPARRNRAPFGQCSCLRPTEDDADPNDSIDPESRETDLSGMDGKSTQVK  
GWLTKQQQVALLIKRLLIARRSKGFFAQIVLPVAVFVCIALVSLVPPFGKYSLE  
LQPMYNEQTVFVSDNAPEDTGTLELNLATKDPGFCRCMEGNPILPDPQAGEEEM  
TTAPVPTIMDLFONGNWTMPNSPACQSSDKIKMLPVCPPGAGGLPQKQNTA  
DILQDLTGRNISDLYLVYQIIISLKNKIVNPEFYGGFSLGVSNTOALPPQSVN  
DAIKQMKHLKAKDSSADRFNLNLRPMTGLDTKNKVNKNKGNHAISSFLAVIN  
NAILRANLQGNPSHYGITAFLNPLNLTKQQLSEVALMTTSDVLVSLVCYFAMSFV  
PASVVFVFLQERYSKAKHLQFSGVFPYIYWLNSFNWCMYVVPATLVIIIFCQQ  
KVSYSNPLPVALALLXGWSITPLMYPASFVKIPSTAYVYVLTNSVNLFIGINGSVA  
TFVLELTQNLKNNIINDILKSVLELPHFCLGRLIDMVKNQAMADALERFINGENRVS  
PLSMDLVGNLFAMAVEGVVFLITVLOVYRFRPRPNKLSPLNDEDEDVRRERQ  
RILDGGQNDILEIKELTKYRBRKRPVADRICVGPCEGCLLVNGVAGKSSFTKM  
LTGDTTVTRGDAFLNKNLSLNHEVHQNMGYCPQFDATITELLTGREHVEFALLRGV  
PEKEVGKVGWAIRKIGLVYGEKYAGNYSNGNKRKLSTAMALMGPPVYVFLDEPTTG  
MDPKARFLWNCALSVYKGRSVYLTSHSMECEALCTRMALMNGRFRCLGSGVQHLK  
NRFGDGYTIWIRIAGSNPDLKPVQDFGLPGLGKVRHNNMLOQLPSSLSLARI  
FSLDSQSKRLHTEDYSVSTQLDQVFNFAKQDQSDDDLKDLSLHKNQTVVDVAULT  
SFLQDKVRESYV"

BASE COUNT 1724 a 1643 c 1759 g 1660 t  
ORIGIN  
Query Match 16.8%; Score 60; DB 9; Length 6786;  
Best Local Similarity 100.0%; Pred. No. 1.9e-23;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 199 ttttgccctcagctgaggttgcgtctgtggaagaccctcactttcagaagaagacaaaca 258  
|||||  
Db 7 TGTGGCCTCAGCTGAGGTGCTGCTGTGGAAGAACCTCACTTTTCAGAAGAAGACAAACA 66  
RESULT 25  
AX253452  
LOCUS Sequence 3 from Patent WO0170810.  
DEFINITION  
ACCESSION AX253452  
VERSION AX253452.1 GI:16073979  
KEYWORDS

SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
TITLE 1 (bases 1 to 7260)  
Schmitz, G. and Bodzioch, M.  
JOURNAL Atp binding cassette transporter 1 (abcl) gene polymorphisms and  
uses thereof for the diagnosis and treatment of lipid,  
cardiovascular or inflammatory disorders  
Patent: WO 0170810-A 3 27-SEP-2001;  
Bayer Aktiengesellschaft (DE)  
FEATURES Location/Qualifiers  
source 1..7260  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 1834 a 1765 c 1905 g 1756 t  
ORIGIN  
Query Match 16.8%; Score 60; DB 6; Length 7260;  
Best Local Similarity 100.0%; Pred. No. 1.9e-23;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 199 tgttgccctcagctgaggttgctgctgtggaagaaccttcacacttcagagaagacaaca 258  
|||||  
Db 327 TGTTCGCTCAGCTGAGGTGCTGCTGTGGAGAACCTCCTCTTCAGAGAAGACAAACA 386  
|||||  
RESULT 26  
AX092594  
LOCUS 7860 bp DNA linear PAT 21-MAR-2001  
DEFINITION Sequence 6 from Patent WO0115676.  
ACCESSION AX092594  
VERSION AX092594.1 GI:13444651  
KEYWORDS human.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
TITLE 1 (bases 1 to 7860)  
Hayden, M.R., Brooks-Wilson, A.R., Pimstone, S.N. and Clee, S.M.  
JOURNAL Compositions and methods for modulating hdl cholesterol and  
triglyceride levels  
Patent: WO 0115676-A 6 08-MAR-2001;  
University of British Columbia (CA) ; Xenon Genetics Inc. (CA)  
FEATURES Location/Qualifiers  
source 1..7860  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 2014 a 1860 c 2008 g 1978 t  
ORIGIN  
Query Match 16.8%; Score 60; DB 6; Length 7860;  
Best Local Similarity 100.0%; Pred. No. 1.9e-23;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 199 tgttgccctcagctgaggttgctgctgtggaagaaccttcacacttcagagaagacaaca 258  
|||||  
Db 81 TGTTCGCTCAGCTGAGGTGCTGCTGTGGAGAACCTCCTCTTCAGAGAAGACAAACA 140  
|||||  
RESULT 27  
AX135712  
LOCUS 7862 bp DNA linear PAT 29-MAY-2001  
DEFINITION Sequence 1 from Patent WO0132184.  
ACCESSION AX135712  
VERSION AX135712.1 GI:14271961  
KEYWORDS human.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 7862)  
AUTHORS Attie, A.D., Cook, M., Gray-Keller, M.P., Hayden, M.R., Pimstone, S. and  
Brooks-Wilson, A.  
TITLE Abcl modulation for the modulation of cholesterol transport  
JOURNAL Patent: WO 0132184-A 1 10-MAY-2001;  
WISCONSIN ALUMNI RESEARCH FOUNDATION (US)  
FEATURES Location/Qualifiers  
source 1..7862  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 2013 a 1861 c 2010 g 1978 t  
ORIGIN  
Query Match 16.8%; Score 60; DB 6; Length 7862;  
Best Local Similarity 100.0%; Pred. No. 1.9e-23;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 199 tgttgccctcagctgaggttgctgctgtggaagaaccttcacacttcagagaagacaaca 258  
|||||  
Db 81 TGTTCGCTCAGCTGAGGTGCTGCTGTGGAGAACCTCCTCTTCAGAGAAGACAAACA 140  
|||||  
RESULT 28  
AX127830  
LOCUS 9741 bp DNA linear PAT 15-MAY-2001  
DEFINITION Sequence 69 from Patent WO0130848.  
ACCESSION AX127830  
VERSION AX127830.1 GI:14134477  
KEYWORDS human.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
TITLE 1 (bases 1 to 9741)  
Deneffe, P., Rosier-Montus, M.F., Arnould-Reguigne, I., Prades, C.,  
Remaley, A., Lemoine, C., Duverger, N., Jaye, M., Searfoss, G.H.,  
Nucleic acids of the human abcl gene and their therapeutic and  
diagnostic application  
Patent: WO 0130848-A 69 03-MAY-2001;  
Aventis Pharma S.A. (FR)  
FEATURES Location/Qualifiers  
source 1..9741  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 2650 a 2180 c 2290 g 2620 t 1 others  
ORIGIN  
Query Match 16.8%; Score 60; DB 6; Length 9741;  
Best Local Similarity 100.0%; Pred. No. 1.9e-23;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 199 tgttgccctcagctgaggttgctgctgtggaagaaccttcacacttcagagaagacaaca 258  
|||||  
Db 191 TGTTCGCTCAGCTGAGGTGCTGCTGTGGAGAACCTCCTCTTCAGAGAAGACAAACA 250  
|||||  
RESULT 29  
AX139817  
LOCUS 9741 bp DNA linear PAT 30-MAY-2001  
DEFINITION Sequence 69 from Patent EP1096012.  
ACCESSION AX139817  
VERSION AX139817.1 GI:14275399  
KEYWORDS human.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
TITLE 1 (bases 1 to 9741)  
Deneffe, P., Rosier-Montus, M.F., Arnould-Reguigne, I., Prades, C.,  
Naudin, L., Lemoine, C., Duverger, N., Jaye, M., Searfoss, I.I., G.H.,



Fri Sep 20 '08:04:06 2002

us-09-846-456-2.oli.rge

Remaley, A., Brewer, H.B. and Dean, M.  
Nucleic acids of the human abcl gene and their therapeutic and  
diagnostic application  
JOURNAL Patent: EP 1096012-A 69 02-MAY-2001;  
Aventis Pharma S.A. (FR)  
FEATURES Location/Qualifiers  
source 1..9741  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 2650 a 2180 c 2290 g 2620 t 1 others  
ORIGIN

Query Match 16.8%; Score 60; DB 6; Length 9741;  
Best Local Similarity 100.0%; Pred. No. 1.9e-23;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 tttggcctcagctgaggttgcctgctgtggaagaacctcactttcagaagaagacaaca 258  
|||||  
Db 191 TGTGGCCTCAGCTGAGGTGCTGCTGTGGAAGAACCTCACTTTTCAGAAGAAGACAACA 250  
|||||

RESULT 30  
AX351038 9741 bp DNA linear PAT 06-FEB-2002  
LOCUS Sequence 10 from Patent WO0183746.  
DEFINITION AX351038  
ACCESSION AX351038  
VERSION AX351038.1 GI:18616393  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 9741)  
AUTHORS Rosier-Montus, M.F., Prades, C., Lemoine, C., Naudin, L., Deneffe, P.,  
Brewer, B., Duverger, N., Remaley, A. and Santamarina-Fojo, S.  
TITLE Regulatory nucleic acid sequences of the abcl gene  
JOURNAL Patent: WO 0183746-A 10 08-NOV-2001;  
Aventis Pharma S.A. (FR)  
FEATURES Location/Qualifiers  
source 1..9741  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 2650 a 2180 c 2290 g 2620 t 1 others  
ORIGIN

Query Match 16.8%; Score 60; DB 6; Length 9741;  
Best Local Similarity 100.0%; Pred. No. 1.9e-23;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 tttggcctcagctgaggttgcctgctgtggaagaacctcactttcagaagaagacaaca 258  
|||||  
Db 191 TGTGGCCTCAGCTGAGGTGCTGCTGTGGAAGAACCTCACTTTTCAGAAGAAGACAACA 250  
|||||

RESULT 31  
AX127831 9854 bp DNA linear PAT 15-MAY-2001  
LOCUS Sequence 70 from Patent WO0130848.  
DEFINITION AX127831  
ACCESSION AX127831  
VERSION AX127831.1 GI:14134478  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 9854)  
AUTHORS Deneffe, P., Rosier-Montus, M.F., Arnould-Reguigne, I., Prades, C.,  
Naudin, L., Lemoine, C., Duverger, N., Jaye, M., Searfoss, G.H.,  
Remaley, A., Brewer, H.B. and Dean, M.  
TITLE Nucleic acids of the human abcl gene and their therapeutic and  
diagnostic application

JOURNAL Patent: WO 0130848-A 70 03-MAY-2001;  
Aventis Pharma S.A. (FR)  
FEATURES Location/Qualifiers  
source 1..9854  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 2665 a 2219 c 2334 g 2635 t 1 others  
ORIGIN

Query Match 16.8%; Score 60; DB 6; Length 9854;  
Best Local Similarity 100.0%; Pred. No. 1.9e-23;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 tttggcctcagctgaggttgcctgctgtggaagaacctcactttcagaagaagacaaca 258  
|||||  
Db 304 TGTGGCCTCAGCTGAGGTGCTGCTGTGGAAGAACCTCACTTTTCAGAAGAAGACAACA 363  
|||||

RESULT 32  
AX139818 9854 bp DNA linear PAT 30-MAY-2001  
LOCUS Sequence 70 from Patent EP1096012.  
DEFINITION AX139818  
ACCESSION AX139818  
VERSION AX139818.1 GI:14275400  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 9854)  
AUTHORS Deneffe, P., Rosier-Montus, M.F., Arnould-Reguigne, I., Prades, C.,  
Naudin, L., Lemoine, C., Duverger, N., Jaye, M., Searfoss, G.H.,  
Remaley, A., Brewer, H.B. and Dean, M.  
TITLE Nucleic acids of the human abcl gene and their therapeutic and  
diagnostic application  
JOURNAL Patent: EP 1096012-A 70 02-MAY-2001;  
Aventis Pharma S.A. (FR)  
FEATURES Location/Qualifiers  
source 1..9854  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 2665 a 2219 c 2334 g 2635 t 1 others  
ORIGIN

Query Match 16.8%; Score 60; DB 6; Length 9854;  
Best Local Similarity 100.0%; Pred. No. 1.9e-23;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 tttggcctcagctgaggttgcctgctgtggaagaacctcactttcagaagaagacaaca 258  
|||||  
Db 304 TGTGGCCTCAGCTGAGGTGCTGCTGTGGAAGAACCTCACTTTTCAGAAGAAGACAACA 363  
|||||

RESULT 33  
AK024328 1556 bp mRNA linear PRI 29-SEP-2000  
LOCUS Homo sapiens cDNA FLJ14266 fis, clone PLACE1002437, highly similar  
DEFINITION to ATP-BINDING CASSETTE TRANSPORTER 1.  
ACCESSION AK024328  
VERSION AK024328.1 GI:10436685  
KEYWORDS oligo capping; fis (full insert sequence).  
SOURCE Homo sapiens placenta cDNA to mRNA, clone PLACE1  
clone: PLACE1002437.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (sites)  
AUTHORS Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, Y.,  
Nishikawa, T., Nagai, K., Sugano, S., Takahashi-Fujii, A., Hara, H.,  
Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K.,  
Arita, M., Nabekura, T., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J.,

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Wakamatsu, A., Nakamura, Y., Nagahari, K., Masuho, Y. and Oshima, A.  
NEDO human cDNA sequencing project  
Unpublished (2000)  
2 (bases 1 to 1556)  
Isogai, T. and Otsuki, T.  
Direct Submission  
Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute,  
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3951, Fax: 81-438-52-3952)  
NEDO human cDNA sequencing project supported by Ministry of  
International Trade and Industry of Japan; cDNA full insert  
sequencing; Research Association for Biotechnology; cDNA library  
construction; 5'- & 3'-end one pass sequencing and clone selection;  
Helix Research Institute (supported by Japan Key Technology Center  
etc.) and Department of Virology, Institute of Medical Science,  
University of Tokyo.

FEATURES  
source  
1..1556  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="PLACE1002437"  
/tissue\_type="placenta"  
/clone\_lib="PLACE1"  
/note="cloning vector: pME18FL3"  
314..1405  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="BAB14887.1"  
/db\_xref="GI:10436686"  
/translation="NACWPLELLKLLNLTERRQCQLLELVAMPFLIFILISVRL  
SYPPYEOHECHPNKAMPAGTLPWQGLIANNPCRYPTPGAPGVGNFNKSIY  
ARLESDARLLYSDIKMDKMRVLRITLQOIKKSSNKLQDFLVNFTSGFLYH  
NLSLPKSTVDKMLRADVLHKVFLQGYQLHLTSLCNGSKSEMIQLGQDEVSELGCLP  
KEKLAARVLRSLMDLPILMDVACDDIAHQLTVPRAAATGDAKPNMMGRET  
LLSICASVPKVFHFHRIHEFCVCSVSLFPAKGVISFWSAFRIWLWKAVFWQ  
HGSEMAVWEGQLGLNTAFETSIDVG"  
BASE COUNT 380 a 363 c 399 g 414 t  
ORIGIN

Query Match 14.3%; Score 51; DB 9; Length 1556;  
Best Local Similarity 100.0%; Pred. No. 3.3e-18;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 208 cagctagcttactgtggaacacctcactttcagaagaagacaaca 258  
|||||  
Db 329 CAGCTGAGTGTGCTGTGGAAGAACCTCACTTTCAGAAGAAGACAACA 379

RESULT 34  
AX092843  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
1..37  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="Synthetic Primer"  
4 a 17 c 11 g 5 t

AX092843  
Sequence 255 from Patent WO0115676.  
AX092843  
AX092843.1 GI:13444900  
synthetic construct.  
synthetic construct.  
artificial sequence.  
1 (bases 1 to 37)  
Hayden, M.R., Brooks-Wilson, A.R., Pimstone, S.N. and Clee, S.M.  
Compositions and methods for modulating hdl cholesterol and  
triglyceride levels  
Patent: WO 0115676-A 255 08-MAR-2001;  
University of British Columbia (CA); Xenon Genetics Inc. (CA)

Query Match 9.8%; Score 35; DB 6; Length 37;  
Best Local Similarity 100.0%; Pred. No. 6.3e-09;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 138 gccgcgtcccttcagggtcccgagccacacagctg 172  
|||||  
Db 1 GCGCGTCCCTTCAGGGCTCCCGAGCCACACGCTG 35

RESULT 35  
AX092841  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
1..38  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="Synthetic Primer"  
6 a 15 c 5 g 12 t

AX092841  
Sequence 253 from Patent WO0115676.  
AX092841  
AX092841.1 GI:13444898  
synthetic construct.  
synthetic construct.  
artificial sequence.  
1 (bases 1 to 38)  
Hayden, M.R., Brooks-Wilson, A.R., Pimstone, S.N. and Clee, S.M.  
Compositions and methods for modulating hdl cholesterol and  
triglyceride levels  
Patent: WO 0115676-A 253 08-MAR-2001;  
University of British Columbia (CA); Xenon Genetics Inc. (CA)

Query Match 9.8%; Score 35; DB 6; Length 38;  
Best Local Similarity 100.0%; Pred. No. 6.3e-09;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 78 cccctcgtcttattcttcagtgtaagaccagcc 112  
|||||  
Db 1 CCCCTCCTGCTTATCTTCAGTTAATGACGACC 35

RESULT 36  
AC091466/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS

AC091466  
Mus musculus clone RP23-353G1, WORKING DRAFT SEQUENCE, 10 unordered  
pieces.  
AC091466  
AC091466.2 GI:14336552  
HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
house mouse.  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
Unpublished  
2 (bases 1 to 207659)  
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,  
Barra, N., Bastien, V., Boguslavsky, L., Boukhaltier, B., Brown, A.,  
Cammarata, J., Campopiano, A., Chang, J., Choepel, Y., Colangelo, M.,  
Collins, S., Collymore, A., Cooke, P., DeAtrellano, K., Dewar, K.,  
Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D.,  
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,  
Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W.,  
Iliev, I., Johnson, R., Jones, C., Karatas, A., LaRocque, K.,  
Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,  
MacLean, C., Macdonald, P., Marquis, N., Matthews, C., McCarthy, M.,  
McEwan, P., McKernan, K., McPheeters, R., Meldrim, J., Meneus, L.,  
Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C.,

Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seanan, S., Severy, P., Sougne, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

# Direct Submission

Submitted (21-APR-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jun 11, 2001 this sequence version replaced gi:13702847.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www.seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L13324

Center clone name: 353\_G.1

----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 203563 bases at least Q40

Consensus quality: 205378 bases at least Q30

Consensus quality: 206185 bases at least Q20

Insert size: 205000; agarose-fp

Insert size: 206759; sum-of-contigs

Quality coverage: 9.6 in Q20 bases; agarose-fp

Quality coverage: 9.5 in Q20 bases; sum-of-contigs

\* NOTE: This is a 'working draft' sequence. It currently consists of 10 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

\* 1 1430: contig of 1430 bp in length  
 \* 1431 1530: gap of 100 bp  
 \* 1531 4004: contig of 2474 bp in length  
 \* 4005 4104: gap of 100 bp  
 \* 4105 132075: contig of 127971 bp in length  
 \* 132076 132175: gap of 100 bp  
 \* 132176 135638: contig of 3463 bp in length  
 \* 135639 135738: gap of 100 bp  
 \* 135739 141621: contig of 5883 bp in length  
 \* 141622 141721: gap of 100 bp  
 \* 141722 157807: contig of 16086 bp in length  
 \* 157808 157907: gap of 100 bp  
 \* 157908 173725: contig of 15818 bp in length  
 \* 173726 173825: gap of 100 bp  
 \* 173826 184364: contig of 10539 bp in length  
 \* 184365 184464: gap of 100 bp  
 \* 184465 206183: contig of 21719 bp in length  
 \* 206184 206283: gap of 100 bp  
 \* 206284 207659: contig of 1376 bp in length.

Location/Qualifiers

1. .207659

/organism="Mus musculus"

/db\_xref="taxon:10090"

/clone="RP23-353G1"

/clone\_lib="RPCI-23 Female Mouse BAC"

1. .1430

misc\_feature

/note="assembly\_fragment"

clone\_end:SP6

vector\_side:left"

misc\_feature 1531. 4004  
 /note="assembly\_fragment"  
 misc\_feature 4105. 132075  
 /note="assembly\_fragment"  
 misc\_feature 132176. 135638  
 /note="assembly\_fragment"  
 misc\_feature 135739. 141621  
 /note="assembly\_fragment"  
 misc\_feature 141722. 157807  
 /note="assembly\_fragment"  
 misc\_feature 157908. 173725  
 /note="assembly\_fragment"  
 misc\_feature 173826. 184364  
 /note="assembly\_fragment"  
 misc\_feature 184465. 206183  
 /note="assembly\_fragment"  
 misc\_feature 206284. 207659  
 /note="assembly\_fragment"

BASE COUNT 58917 a 43251 c 44962 g 59609 t 920 others  
 ORIGIN  
 vector\_side:right"

Query Match 7.6%; Score 27; DB 2; Length 207659;  
 Best Local Similarity 100.0%; Pred. No. 0.00044;  
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 ttatttttcagtaataatgacccac 114

|||||  
 Db 175462 TTTATCTTTCAGTTAATGACCCAC 175436

RESULT 37

AF287263

LOCUS

DEFINITION

AF287263

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

mrna

1. .278572

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="BAC 129K10: 142M4; 197F14"

join(87184. .87402,103551.103699,115064. .115187,

119217. .119358,120364. .120482,133479. .133600,

135582. .135758,147083. .147175,149791. .150031,

152367. .152506,152845. .152961,156570. .156767,

157459. .157664,158200. .158376,158821. .159043,

160617. .160838,161790. .161994,163076. .163189,

164646. .164817,165738. .165869,167171. .167313,

167492. .167629,168320. .168540,169790. .169862,

170735. .170937,172636. .172684,172855. .172968,

173925. .174073,175605. .175729,176761. .176859,

176859. .176859,176761. .176859,

176859. .176859,176761. .176859,

176859. .176859,176761. .176859,

176859. .176859,176761. .176859,

176859. .176859,176761. .176859,

176859. .176859,176761. .176859,

176859. .176859,176761. .176859,

176859. .176859,176761. .176859,

176859. .176859,176761. .176859,

176859. .176859,176761. .176859,

176859. .176859,176761. .176859,

```

180181. 180370,181179. 181273,182395. 182427,
183398. 183503,191761. 191829,192306. 192475,
193631. 193808,196619. 196734,196845. 196989,
198526. 198649,200922. 201051,201281. 201401,
202114. 202176,202890. 202996,204955. 205096,
205427. 205558,206739. 206842,207361. 207453,
208135. 208378,209638. 210783)
/gene="Abcal"
/product="ATP-binding cassette 1, sub-family A, member 1"
87184. 210783
/gene="Abcal"
/join(119237..119358,120364..120482,133479..133600,
135382..135738,147083..147175,149791..150031,
152367..152506,152845..152961,156570..156767,
175459..157664,158200..158376,158821..159043,
160617..160838,161790..161994,163076..163189,
164646..164817,165738..165869,167171..167313,
167492..167629,168320..168540,169790..169862,
170735..170937,172636..172684,172855..172968,
173925..174073,175605..175729,176761..176859,
180181..180370,181179..181273,182395..182427,
183398..183503,191761..191829,192306..192475,
193631..193808,196619..196734,196845..196989,
198526..198649,200922..201051,201281..201401,
202114..202176,202890..202996,204955..205096,
205427..205558,206739..206842,207361..207453,
208135..208378,209638..209778)
/gene="Abcal"
/codon_start=1
/product="ATP-binding cassette 1, sub-family A, member 1"
/protein_id="AAG39073.1"
/db_xref="GI:11611825"
/translation="MPSAGTLPWVQGIICNANPCRYPTTGPAGVGVGNFKSIVSR
LFSDAQLLYSDRTSIKDMHVKRLQIKHPNSNKLQDLVDNETFSGFLQHNL
SPRSTVLIQANGLQGVFLQYOLHSLASGLQKLEIEIOLGDAEYSAQLGLPRK
KLDAAEVRLNMDILKPVYTKLNSTLPTOHLARATVLLSLGLAQLAELSTSK
SDMRQVMTNVNSSSSQIYQAVSRIVCGHPGGGLKIKSLNWDNNKALFGG
NNTEDVDFYDNSTPYCNDLKNLESSPLRIKALKAPLLVGLKILYTPDPATRG
VMAEVKTFGLAVFHLDECMWELSPQIWFNEMSEMDLVRLDLSRGNQDFWQK
LGLDWTADQIMFLAKNPEDVQSPGVSTVREAFNETNOAIQIISFMECVNLKL
EPITPEVRLINKSMELDERKFWAGVFTGIDPSVELPHVHYKIRKMDINVERTNK
IKQGVDFRMPADFPEDRYWGVGFAYLQDVQEIIRVLTSEKKTGVVQOVPYPC
YVDIFLRVGRSMPLFMTLAWISVAVIISVYKEARLKTETMRIMGLDGLWFS
WFYSLLPLLVAGLLVILKGLNLLPYSDPSVFLSVFVFWMTIQLQFLISTFSR
ANLAAAGGIIYPTLIVPLVLCVQDYVGFSGIKFASLLSVAFGFGCEYFALFEEQ
GIGVQNDLFESEVEDGFLNTAVSMMLDFLYIGVMTWYIEAVPGQYIGTPRPYF
PCTKSWFGEIEDEKSHPSQSQSEICMEEPETHRLGVSTQNIKVYRDMKVAV
DGLALNFEQITSLFHNGAGKTTMSILGLPPTSGTAVILGDKIRSEMSIRON
LGVCPQHNVLFDMLTVEEHIFVYARLGLSEKHVKAEMQMALDVLGPPSKLSTSQ
LSGMORKLKVALAFVGGKSVILDEPTAGVDPSYRGGIWEILLKYRCQRTILSTHH
MDEADILGDRIATIIISHKLCVGSLLFNQJLGTGYLLILVKDVESLSRNSST
VSLCKREDVSQSSDAGLGHESDITLIDVSAIINLIRKHVSEARVEDIGHETLY
VLPEAKGAFVELHEIDRDLSDIGISYGETTLEIFLVAESGVDAETSDG
TLFARNRRAFGDKQLHPFTEDDVAVDNDSIDPESRETLLSGMDGKGYQLKGW
KLQQQFVALLKRLIARRSRGFFAIVLPAVFCIALVFSLI VPPGKYPSLELQ
PMNYQYTVSDNFTNGMTQELLNALTQDPGFCRMEGNIPTDPCLAGEDWTI
SPVQSTVDLFGNMTKNSPACOCSSDKIKMLPYCPGAGGLPPQPKQKTADI
LONLGRNLSLXVLYVQIILAKLANKTWNEFRIGGSLGVNSQALPPSHVEDNA
IKQMKLKLLTSADRLSLGFRMAGLTDKNNKVMFNKGMHAISSVFNNAIL
RANLQGENPSQGITAFNHLNLTQQLSEVALMTSDVILVSVICIFAMFVPASV
VFLIQRYSKAKHQFISGVPIVYLSFNVMDCNYYVPATVLIIFICFOQKSYV
SSTNPLVALLLLIGWSTIPLMYPASVFKIPSTAYVLTVSNLFIGVSVATFVL
ELVTNNKNDINDILKSVFLIPHFCGLRGLIDVMKNOAMADALFRGNRVSLSW
DLVGRNLKAMAGGVVFLITVLIQYRFFIRPRPKAPLPLNDEDEDVRRERQLID
GGQNDILEIKELTKIYRKRKPAVDRCIGIPPGECFLGVNGAGKSTFKMLTD
TPVGRDAFLKNSILSNIEVHQNNQCPQFADITELTGREHVEFFALLRGVEKE
VGVGEAIKGLGLVYGERYASNYGNGKRLSTAMALIGVPPVFLDEPTGMDPKA
RFTLVNRIAGSNPKLPVQEFFLAPGSLVKEKHMLOYLPSSLSLARIFSLIS
QSKRLIHIEDYSVQSTLQDVFNFAKQSDDDHLKDLSLKHQIVDVAVLTSLQD
EKVKESV"
76801 a 58762 c 61256 g 81498 t 255 others
BASE COUNT
ORIGIN

```

```

Query Match          7.6%; Score 27; DB 10; Length 278572;
Best Local Similarity 100.0%; Pred. No. 0.00044;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 ttatctttcagtttaagaccagccac 114
|||||
Db 103522 TTTATCTTTCAGTTTAAAGCAGCCAC 103548
|||||

```

```

RESULT 38
AX239607
LOCUS AX239607 12425 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 50 from Patent WO0164874.
ACCESSION AX239607
VERSION AX239607.1 GI:15797282
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 12425)
AUTHORS Lees,A.M., Lees,R.S., Law,S.W. and Arjona,A.A.
TITLE Low density liprotein binding proteins and their use in diagnosing
and treating atherosclerosis
JOURNAL Patent: WO 0164874-A 50 07-SEP-2001;
Boston Heart Foundation, Inc. (US)
FEATURES
source
location/Qualifiers
1..12425
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2563 a 3702 c 3566 g 2594 t
ORIGIN

```

```

Query Match          6.7%; Score 24; DB 6; Length 12425;
Best Local Similarity 100.0%; Pred. No. 0.022;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 ggaggaggaggaggaggaggagg 349
|||||
Db 2395 GGAGGAGGAGGAGGAGGAGGAAG 2418
|||||

```

```

RESULT 39
AC006326/c
LOCUS AC006326 18427 bp DNA linear PRI 21-DEC-1999
DEFINITION Homo sapiens BAC clone GSI-234B20 from 7q31.1-q31.3, complete
sequence.
ACCESSION AC006326
VERSION AC006326.2 GI:4508133
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 18427)
AUTHORS Sulston,J.E. and Waterston,R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
REFERENCE 2 (bases 1 to 18427)
AUTHORS Maupin,R., Bauer,C. and Le,T.
TITLE The sequence of Homo sapiens BAC clone GSI-234B20
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 18427)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (09-JAN-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4 (bases 1 to 18427)

```

AUTHORS TITLE JOURNAL	REFERENCE AUTHORS TITLE JOURNAL	REFERENCE AUTHORS TITLE JOURNAL	REFERENCE AUTHORS TITLE JOURNAL	COMMENT
-----------------------------	--	--	--	---------

Waterston, R.H.  
Direct Submission  
Submitted (24-MAR-1999) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
5 (bases 1 to 18427)  
Waterston, R.  
Direct Submission  
Submitted (01-MAY-1999) Department of Genetics, Washington  
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
6 (bases 1 to 18427)  
Waterston, R.  
Direct Submission  
Submitted (10-JUL-1999) Department of Genetics, Washington  
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
7 (bases 1 to 18427)  
Waterston, R.  
Direct Submission  
Submitted (21-DEC-1999) Department of Genetics, Washington  
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
On Mar 24, 1999 this sequence version replaced gi:4138782.  
----- Genome Center  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: <http://genome.wustl.edu/gsc>  
Contact: [sapiens@watson.wustl.edu](mailto:sapiens@watson.wustl.edu)  
----- Summary Statistics  
-----  
Center project name: H GS234B20

**NOTICE:** This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

## MAPPING INFORMATION:

**MAPPING INFORMATION:**  
The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/STB/CH07>, send <mailto:ericgreen@nih.gov>, or see <http://genome.wustl.edu/gsc>

**SOURCE INFORMATION:**

SOURCE INFORMATION:  
This clone is from the first BAC library from Genome Systems, Inc.  
(<http://www.genomesystems.com>).  
cell line: lymphoblastoid

Cell line: Lymphoblastoid

Haplotypes: two

HAPLOYPES: LWO  
 VECTOR: pBELoBAC

VECTOR: pBEL0BAC

Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:  
The clone sequenced to the left is RP5-866N18, 200 bp overlap;  
the clone sequenced to the right is CTA-343P13, 200 bp overlap. The  
actual start is unknown, the first known base of overlap is at base  
position 139224 of RG11A06; actual end is at 53881 of RA-343P13.

Location/Qualifiers

Location,
1 1842

1. 18427

/organism="Homo sapien

/db\_xref="taxon:960

```
/chromosome="7"
```

```

/chromosome="/"
/pan="7a31 1-a31 3"

```

```
/map="7q31.1-q31.3"
```

FEATURES

repeat_region	75. .119 /rpt_family="AT_rich"
repeat_region	436. .707 /rpt_family="Other"
repeat_region	708. .947 /rpt_family="L1"
misc_feature	/note="HNENP A1 pseudogene"
repeat_region	1497. .1613 complement(1374. .3465)
repeat_region	3523. .3646 /rpt_family="(TGG)n"
repeat_region	3673. .3965 /rpt_family="(GGCA)n"
repeat_region	4000. .4028 /rpt_family="Alu"
repeat_region	4029. .4162 /rpt_family="(GAAAA)n"
repeat_region	4381. .4674 /rpt_family="Alu"
repeat_region	5298. .5473 /rpt_family="MER86"
repeat_region	6379. .6740 /rpt_family="MaLR"
repeat_region	7026. .7046 /rpt_family="AT_rich"
repeat_region	7246. .7544 /rpt_family="Alu"
repeat_region	7640. .7666 /rpt_family="AT_rich"
repeat_region	8042. .8132 /rpt_family="L2"
repeat_region	9524. .9818 /rpt_family="Alu"
repeat_region	10661. .10660 /rpt_family="MER1_type"
repeat_region	10955. .11122 /rpt_family="Alu"
repeat_region	11123. .11142 /rpt_family="(TAAAA)n"
repeat_region	11143. .11281 /rpt_family="Alu"
misc_feature	12018. .12343 /note="match to EST AA532765 (NID:g2278341) nf70b01.sl"
misc_feature	12068. .12557 /note="Similar to Mus musculus EST AA220774 (NID:gl838594) mv69b03.rl"
repeat_region	13097. .13385 /rpt_family="Alu"
misc_feature	13911. .14272 /note="match to EST AI033200 (NID:g3254153) ow94f08.sl"
misc_feature	14430. .14842 /note="match to EST T95217 (NID:g733841) ye44e03.rl"
misc_feature	15142. .15277 /note="match to EST AA938985 (NID:g3098898) op75f04.sl"
misc_feature	15143. .15561 /note="match to EST W93113 (NID:g1422275) zh45f06.sl"
misc_feature	15725. .16002 /note="match to EST AI202170 (NID:g3754776) qs69b06.sl"
misc_feature	15751. .16311 /note="match to EST H04382 (NID:g867315) yj21b11.sl"
misc_feature	15761. .16347 /note="match to EST W17194 (NID:g1291638) zbl7a04.rl"
misc_feature	15767. .16191 /note="match to EST W93205 (NID:g1422377) zh45f06.rl"
misc_feature	15768. .16228 /note="match to EST W78828 (NID:g1241529) zbl7a04.sl"
misc_feature	16018. .16276 /note="match to EST AA601910 (NID:g2436063) np02c01.sl"
misc_feature	16019. .16778 /note="match to EST AI379153 (NID:g4189006) tc66a06.sl"
misc_feature	16020. .16285 /note="similar to EST T99055 (NID:g748792) ye67a02.sl"

```

misc_feature 16021..16482
/note="match to EST AI249171 (NID:g3844568) qh68f12.x1"
misc_feature 16059..16399
/note="match to EST AA676311 (NID:g2656833) z110g04.s1"
misc_feature 16072..16571
/note="match to EST W88685 (NID:g1404396) zh71e04.r1"
misc_feature 16316..16832
/note="match to EST AI200003 (NID:g3752609) qf90c03.x1"
misc_feature 16517..16896
/note="match to EST W78848 (NID:g1389395) zh51g03.r1"
gene 16364..16891
/complement(16664..>16891)
CDS 16364..16891
/gene="WUGSC:H_GS234B20.1"
/complement(16664..>16891)
/note="WUGSC:H_GS234B20.1"
/codon_start=1
/evidence=not_experimental
/protein_id="AAD28351.1"
/db_xref="GI:4731069"
/translation="SLGTAGRVCMLTRGMDSCVEMCCGRGYDTSHTVTRMTKCGKRFH
MCCAVRCQDCLEALDVHTCKAPKNADWTAT"
BASE COUNT 5121 a 3991 c 3520 g 5795 t
ORIGIN

Query Match
Best Local Similarity 6.7%; Score 24; DB 9; Length 18427;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 ggagggaggaggaggaggaggagg 349
|||||
Db 3652 GCAGGAGGAGGAGGAGGAGGAG 3629

RESULT 40
AL583882
LOCUS Human DNA sequence from clone RP5-1098C18 on Chromosome
DEFINITION lp36-23-36.33, complete sequence.
ACCESSION AL583882
VERSION AL583882.6 GI:13990182
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 90582)
Thomas, D.
Direct Submission
Submitted (03-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
Requests: clonerequests@sanger.ac.uk
On May 7, 2001 this sequence version replaced gi:13897175.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL; Sw:,
SWISSPROT; Tr:, TrEMBL; Wp:, WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/projects/Celegans/wormpep
This sequence
was generated from part of bacterial clone contigs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping

```

Group. Further information can be found at  
<http://www.sanger.ac.uk/HGP/Chrl>  
 RP5-1098C18 is from the library RPCI-5 constructed by the group of  
 Pieter de Jong. For further details see  
<http://www.chori.org/bacpac/home.htm>  
 VECTOR: pCYPAC2  
 This sequence is the entire insert of clone RP5-1098C18 The true  
 right end of clone RP11-154H17 is at 1402 in this sequence.

```

FEATURES
    source
        Location/Qualifiers
            1..90582
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /chromosome="1"
                /map="p36.23-36.33"
                /clone="RP5-1098C18"
                /clone_lib="RPCI-5"
                635..666
                /note="16 copies 2 mer tg 100% conserved"
                954..1083
                /note="13 copies 10 mer gtgtgtgtgt 66% conserved"
                982..1079
                /note="49 copies 2 mer gt 72% conserved"
                984..1079
                /note="24 copies 4 mer gtgt 72% conserved"
                1183..1286
                /note="MIR repeat: matches 41..150 of consensus"
                2755..2968
                /note="MT1B repeat: matches 1..223 of consensus"
                complement(2762..3096)
                /note="match: STS: Em:HSPE03C10"
                3462..3740
                /note="L1PB2 repeat: matches 5861..6155 of consensus"
                3890..3991
                /note="MER20 repeat: matches 118..218 of consensus"
                4018..4076
                /note="MER20 repeat: matches 1..52 of consensus"
                4909..4961
                /note="12 repeat: matches 2693..2745 of consensus"
                5551..5610
                /note="6 copies 10 mer tacacacaca 76% conserved"
                5552..5607
                /note="14 copies 4 mer acac 75% conserved"
                5553..5610
                /note="29 copies 2 mer ca 75% conserved"
                6314..6690
                /note="L1MC3 repeat: matches 5437..5825 of consensus"
                6702..6989
                /note="AluDb repeat: matches 13..298 of consensus"
                6992..7297
                /note="AluSg repeat: matches 1..306 of consensus"
                7301..7532
                /note="AluJ repeat: matches 63..291 of consensus"
                7540..8738
                /note="L1MC3 repeat: matches 5843..7037 of consensus"
                8740..8801
                /note="Alu repeat: matches 242..298 of consensus"
                8864..9575
                /note="L1MC3 repeat: matches 7021..7738 of consensus"
                9646..10004
                /note="MT1A1 repeat: matches 1..363 of consensus"
                10015..10187
                /note="AluSp repeat: matches 132..304 of consensus"
                10338..10458
                /note="AluSp repeat: matches 1..121 of consensus"
                11856..11944
                /note="L1M4 repeat: matches 3850..3937 of consensus"
                11945..12188
                /note="MER8 repeat: matches 2..238 of consensus"
                12189..12628
                /note="L1M4 repeat: matches 3937..4387 of consensus"
                12629..12929
                /note="AluSg repeat: matches 1..301 of consensus"
                12934..13124

```

```

/note="LIMA9 repeat: matches 6028. .6211 of consensus"
13050. .13154
/note="LIMA9 repeat: matches 6199. .6303 of consensus"
13160. .13423
/note="LIMA4 repeat: matches 4374. .4642 of consensus"
13425. .13600
/note="AluSg/x repeat: matches 130. .305 of consensus"
13601. .13659
/note="LIMA4 repeat: matches 4648. .4709 of consensus"
13660. .13966
/note="AluSg repeat: matches 1. .312 of consensus"
13976. .14147
/note="43 copies 4 mer gga 70% conserved"
13984. .14151
/note="6 copies 28 mer 72% conserved"
14026. .14125
/note="10 copies 10 mer gagagaaga 72% conserved"
14149. .14697
/note="LIMA4 repeat: matches 4694. .5255 of consensus"
14924. .14956
/note="LIMB3 repeat: matches 6153. .6183 of consensus"
14957. .15266
/note="AluSp repeat: matches 1. .311 of consensus"
15267. .15610
/note="LIMB3 repeat: matches 5792. .6153 of consensus"
15853. .16166
/note="AluSx repeat: matches 1. .311 of consensus"
16201. .16511
/note="AluY repeat: matches 1. .311 of consensus"
17222. .17364
/note="MER5A repeat: matches 20. .161 of consensus"
18757. .19068
/note="MLT1J repeat: matches 8. .330 of consensus"
19575. .19996
/note="MLT1C repeat: matches 1. .466 of consensus"
20098. .20189
/note="23 copies 4 mer gatg 64% conserved"
2285. .23136
/note="9 copies 28 mer 66% conserved"
22893. .23096
/note="51 copies 4 mer gga 66% conserved"
22964. .23123
/note="16 copies 10 mer gaaaggaag 60% conserved"
23620. .23685
/note="MIR repeat: matches 20. .85 of consensus"
24243. .24699
/note="CpG island"
evidence-not_experimental
24326. .24541
/note="4 copies 54 mer 94% conserved"
25213. .25295
/note="L2 repeat: matches 2595. .2682 of consensus"
25770. .25958
/note="3 copies 63 mer 97% conserved"
26810. .27670
/note="L1P11 repeat: matches 5310. .6159 of consensus"
27699. .27824
/note="MIR repeat: matches 81. .212 of consensus"
27909. .28364
/note="MLT1D repeat: matches 17. .505 of consensus"
28426. .28538
/note="MIR repeat: matches 28. .149 of consensus"
28764. .29250
/note="MER74A repeat: matches 1. .558 of consensus"
29282. .29456
/note="MER31A repeat: matches 18. .196 of consensus"
30627. .30905
/note="AluSg repeat: matches 18. .296 of consensus"
32695. .32726
/note="L2 repeat: matches 2466. .2497 of consensus"
33150. .33215
/note="MIR repeat: matches 52. .114 of consensus"
34512. .34640

```

```

repeat_region
/note="FLAM_C repeat: matches 2. .127 of consensus"
34647. .34702
/note="2 copies 28 mer 96% conserved"
34648. .34703
/note="14 copies 4 mer atat 96% conserved"
34703. .34760
/note="FAM repeat: matches 116. .173 of consensus"
34766. .34906
/note="AluSg/x repeat: matches 1. .139 of consensus"
35129. .35212
/note="MER96 repeat: matches 1. .84 of consensus"
35287. .35769
/note="LIMA8 repeat: matches 5812. .6290 of consensus"
35772. .35847
/note="AluSg/q repeat: matches 226. .299 of consensus"
35853. .35924
/note="MLT2B repeat: matches 1. .73 of consensus"

Query Match 6.7%; Score 24; DB 9; Length 90582;
Best Local Similarity 100.0%; Pred. No. 0.024;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 ggaggaggaggaggaggaggagg 349
|||||
DB 60005 GGAGGAGGAGGAGGAGGAGGAG 60028

RESULT 41
AF179296/c
LOCUS AF179296
DEFINITION Homo sapiens, *** SEQUENCING IN PROGRESS ***, in ordered pieces.
ACCESSION AF179296
VERSION AF179296.1 GI:5802962
KEYWORDS HTG; HTGS_PHASE2.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 123339)
AUTHORS Schudy,A., Platzner,M., Schilhabel,M., Koczan,D., Thiesen,H.-J.,
Merck-Rousseau,M.F., Baumgart,C., Menzel,U., Weber,J.,
Schattke,R. and Rosenthal,A.
TITLE Direct Submission
JOURNAL Submitted (20-AUG-1999) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
COMMENT contig 1: pos. 1 - 92151 contig 2: pos. 92152 - 123339 Inbetween a
gap of unknown size.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES
source
1. .123339
/organism="Homo sapiens"
/db_xref="taxon.9606"
complement(2. .144)
/rpt_family="MIR"
395. .438
/note="SST_REPEAT"
complement(556. .676)
/rpt_family="MIR"
566. .636
/rpt_family="mir"
complement(746. .930)
/rpt_family="MIR"
1036. .1142
/rpt_family="alu"
1045. .1119
/rpt_family="sva"
1055. .1182
/rpt_family="alu"
1055. .1180
/rpt_family="ALU"

```

```

repeat_region 1055..1335 /rpt_type=Inverted
repeat_region 1055..1331 /rpt_family="AluSx"
repeat_region complement(1077..1179) /rpt_family="SVA"
repeat_region 1175..1232 /rpt_family="SVA"
repeat_region 1175..1232 /rpt_family="SVA"
repeat_region 1175..1232 /rpt_family="SVA"
repeat_region 1229..1333 /rpt_family="Alu"
repeat_region 1232..1335 /rpt_family="Alu"
repeat_region 1272..1346 /rpt_family="ALU"
repeat_region /note="SST_REPEAT" 1272..1346
repeat_region /note="XNUN_REPEAT" complement(1332..1628)
repeat_region /rpt_family="AluSx" complement(1339..1627)
repeat_region /rpt_family="Alu" complement(1339..1628)
repeat_region /rpt_family="ALU" 1347..1628
repeat_region /rpt_type=Inverted complement(1366..1423)
repeat_region /rpt_family="SVA" complement(1410..1474)
repeat_region /rpt_family="SVA" complement(1412..1488)
repeat_region /rpt_family="Alu" complement(1431..1515)
repeat_region /rpt_family="SVA" complement(1502..1605)
repeat_region /rpt_family="SVA" 1506..1605
repeat_region /rpt_family="SVA" complement(1535..1628)
repeat_region /rpt_family="Alu" complement(1566..1628)
repeat_region /rpt_family="SVA" 1624..1730
repeat_region /rpt_family="MIR" complement(1624..1737)
repeat_region /rpt_family="mir" 1629..1737
repeat_region /rpt_family="MIR" 1796..1841
repeat_region /note="XNUN_REPEAT" 1799..1841
repeat_region /note="SST_REPEAT" complement(1800..1985)
exon /note="GRAIL"
repeat_region /evidence=not_experimental 1966..2073
repeat_region /note="SST_REPEAT" 2756..2847
repeat_region /rpt_family="Alu" 2756..2824
repeat_region /rpt_family="SVA" 2758..3051
repeat_region /rpt_family="AluSx" 2758..2869
repeat_region /rpt_family="Alu" 2777..2872
repeat_region /rpt_family="SVA" complement(2777..2868)
repeat_region /rpt_family="SVA" complement(2822..3037)
exon /note="FXHB"
repeat_region /evidence=not_experimental 2863..3047
/rpt_family="ALU" 2867..3048
/rpt_family="Alu" 2883..2954
/rpt_family="SVA" 2894..2987
/rpt_family="Alu" 2913..2976
/rpt_family="SVA" 2962..3017
/rpt_family="SVA" 3041..3105
/note="SST_REPEAT" 3057..3208
/rpt_family="lmc3" complement(3070..3271)
/rpt_family="lmc3" 3137..3205
/rpt_family="lmc4" 3138..3205
/rpt_family="lmc2" complement(3142..3207)
/rpt_family="lmc3" complement(3371..3421)
/rpt_family="mir" 3372..3545
/rpt_family="MIR" complement(3405..3461)
/rpt_family="mir" complement(3721..3802)
/rpt_family="mir" 3762..3900
/rpt_family="MIR" 4007..4055
/note="SST_REPEAT" 4108..4318
/rpt_family="MSTB" 4108..4246
/rpt_family="MSTA" complement(4110..4245)
/rpt_family="msta" complement(4154..4218)
/rpt_family="thelb" 4211..4246
/note="XPOUND"
/evidence=not_experimental 4290..4414
/rpt_family="Alu" 4319..4619
/rpt_family="AluSx" 4322..4449
/rpt_family="SVA" 4322..4491
/rpt_family="Alu" 4332..4484
/rpt_family="ALU" complement(4342..4443)
/rpt_family="SVA" 4345..4706
/rpt_type=Inverted 4347..4706
/rpt_type=Inverted 4349..4706
/rpt_type=Inverted 4351..4417
/rpt_type=Inverted 4465..4606
/rpt_family="Alu" 4524..4579
/rpt_family="SVA" 4709..5070
/rpt_type=Inverted 4709..5066

```



```
repeat_region /rpt_type=inverted 4709..5068
repeat_region /rpt_type=inverted 4775..5004
repeat_region /note="XNUN_REPEAT" 4775..4866
repeat_region /rpt_type=tandem complement(4775..4866)
repeat_region /note="(TGGA)n"
/rpt_family="Simple_repeat"
```

Query Match 6.7% Score 24; DB 2; Length 123339;  
Best Local Similarity 100.0%; Pred. No. 0.024;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 326 ggagggaggaggaggaggagg 349  
|||||  
Db 82626 GGAGGAGGGAGGGAGGGAGG 82603

RESULT 42  
AC036144/c  
LOCUS AC036144 127270 bp DNA linear PRI 18-JAN-2002  
DEFINITION Homo sapiens chromosome 8, clone RP11-29316, complete sequence.  
ACCESSION AC036144  
VERSION AC036144.7 GI:18104887  
KEYWORDS HTG.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 127270)  
Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
JOURNAL  
Unpublished  
2 (bases 1 to 127270)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,  
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,  
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,  
Collins,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S.,  
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,  
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,  
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,  
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,  
Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J.,  
Levine,R., Liu,G., Locke,K., Macdonald,P., Marquis,N.,  
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,  
Meldrim,J., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,  
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
O'Neill,D., Olivat,T.M., Oliver,J., Peterson,K., Pierre,N.,  
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,  
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,  
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,  
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,  
Young,G., Zainoun,J., Zimner,A. and Zody,M.

Direct Submission  
Submitted (07-APR-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 127270)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,  
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,  
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,  
Choepel,Y., Colangelo,M., Collins,S., Collimore,A., Cooke,A.,  
Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S., Fargo,S.,  
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,  
Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,  
Jones,C., Kamat,A., Karatas,A., Kells,C., Labocque,K.,  
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,  
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,  
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,  
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,  
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,  
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,  
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,  
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R.,  
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,  
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission  
Submitted (10-JAN-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
5 (bases 1 to 127270)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,  
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,  
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,  
Choepel,Y., Colangelo,M., Collins,S., Collimore,A., Cooke,A.,  
Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S., Fargo,S.,  
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,  
Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,  
Jones,C., Kamat,A., Karatas,A., Kells,C., Labocque,K.,  
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,  
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,  
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,  
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,  
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,  
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,  
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,  
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R.,  
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,  
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission  
Submitted (08-OCT-2001) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
4 (bases 1 to 127270)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,  
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,  
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,  
Choepel,Y., Colangelo,M., Collins,S., Collimore,A., Cooke,A.,  
Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S., Fargo,S.,  
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,  
Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,  
Jones,C., Kamat,A., Karatas,A., Kells,C., Labocque,K.,  
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,  
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,  
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,  
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,  
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,  
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,  
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,  
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R.,  
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,  
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission  
Submitted (18-JAN-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jan 10, 2002 this sequence version replaced gi:15983570.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>

Direct Submission  
Submitted (18-JAN-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jan 10, 2002 this sequence version replaced gi:15983570.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>

Contact: sequence\_submissions@genome.wi.mit.edu  
 ----- Project Information  
 Center project name: L8779  
 Center clone name: 293\_L\_6  
 -----

FEATURES  
 source T7 end overlaps AC009941 [WTCGR Project L1987] by 103216 bp.

```

Location/Qualifiers
1. .127270
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8"
/clone="RP11-293L6"
/clone_lib="RPC1-11 Human Male BAC"
complement(4. .3746)
/rpt_family="L1"
946. .953
/note="<30 qual SNGL region"
1029. .1033
/note="<30 qual SNGL region"
1056. .1060
/note="<30 qual SNGL region"
1110. .1126
/note="<30 qual SNGL region"
1135. .1141
/note="<30 qual SNGL region"
1145. .1152
/note="<30 qual SNGL region"
1169. .1177
/note="<30 qual SNGL region"
1190. .1195
/note="<30 qual SNGL region"
1219. .1224
/note="<30 qual SNGL region"
1230. .1234
/note="<30 qual SNGL region"
2097. .2102
/note="<30 qual SNGL region"
complement(3768. .3859)
/rpt_family="L1"
complement(3860. .4173)
/rpt_family="AluY"
complement(4179. .6609)
/rpt_family="L1PA7"
6610. .6647
/rpt_family="(TTG)n"
complement(6648. .8549)
/rpt_family="L1PA7"
complement(8549. .9721)
/rpt_family="L1"
complement(9736. .10077)
/rpt_family="L1"
10180. .10236
/rpt_family="AT_rich"
complement(10474. .10599)
/rpt_family="MIR"
10698. .10800
/rpt_family="MIRNA"
11389. .11416
/rpt_family="AT_rich"
complement(12357. .12484)
/rpt_family="MIR3"
complement(12621. .12742)
/rpt_family="L2"
complement(12759. .13088)
/rpt_family="MER7A"
complement(13108. .13281)
/rpt_family="MER5B"
complement(13285. .13379)
/rpt_family="MER5B"
14034. .14057
/rpt_family="AT_rich"

```

```

repeat_region complement(14859. .14987)
repeat_region /rpt_family="L2"
repeat_region 15347. .15525

Query Match
Best Local Similarity 6.7%; Score 24; DB 9; Length 127270;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 ggaggaggaggaggaggaggagg 349
DB 18870 GGAGGAGGAGGAGGAGGAGGAGG 18847

RESULT 43
AL158043/c
LOCUS
DEFINITION
AL158043
Human DNA sequence from clone RP11-5N23 on chromosome 10p14-15.3
Contains ESTs, GSSs, STSS and a CpG island. Contains the 5' part of
the PRKCQ gene for protein kinase C theta and a novel gene,
complete sequence.
AL158043
AL158043.14 GI:9863558
HTG; CpG island; PRKCQ; protein kinase.
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 132741)
AUTHORS
Chapman,J.
TITLE
Direct Submission
JOURNAL
Submitted (09-OCT-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Aug 21, 2000 this sequence version replaced gi:9856693.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/C.elegans/wormpep/ This sequence
was generated from part of bacterial clone contigs of human
chromosome 10, constructed by the Sanger Centre Chromosome 10
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr10
RP11-5N23 is from the library RPC1-11.1 constructed at the Roswell
Park Cancer Institute by the group of Pieter de Jong. For further
details see http://bacpac.med.buffalo.edu/
VECTOR: pBACE3.6
IMPORTANT: This sequence is not the entire insert of clone
RP11-5N23 It may be shorter because we sequence overlapping
sections only once, except for a 100 base overlap.
The true right end of clone RP11-5N23 is at 132741 in this
sequence. The true left end of clone RP11-55418 is at 102011 in
this sequence. The true right end of clone RP11-563J2 is at 100 in
this sequence.

```

```

FEATURES
source Location/Qualifiers
1. .132741
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="10"
/map="p14-15.3"
/clone="RP11-5N23"

```

```

/clone_lib="RPCI-11.1"
complement(join(<510. .619,6158. .6287,11510. .11595,
11695. .11726,12871. .13033,21911. .21971,25470. .25669,
29493. .29610,94686. .94767))
/gene="PRKCQ"
/note="match: cDNAs: Em:L01087 Em:222521 Em:D11091
Em:L07032 Em:L07860 Em:L07861 Em:X60304 Em:AB011812
Em:AF121936 Em:L07830 Em:AF251036 Em:M69042 Em:D10495
match: ESTs: Em:A152852 Em:AA153683 Em:A1326731"
/product="ba5N23.1 (protein kinase C theta)"
/evidence=not_experimental
complement(510. .94767)
/gene="PRKCQ"
complement(join(<510. .619,6158. .6287,11510. .11595,
11695. .11726,12871. .13033,21911. .21971,25470. .25669,
29493. .29610))
/gene="PRKCQ"
/note="continued in ba563J2 (AL137145), gene ba563J2.2
match: proteins: Sw:Q024759 Sw:Q02111 Tr:Q92333 Sw:P09215
Sw:Q05655 Sw:P28867 Tr:Q15144 Sw:P34722 Wp:B0545B
Tr:O61225 Tr:O62569"
/codon_start=1
/evidence=not_experimental
/product="ba5N23.1 (protein kinase C theta)"
/protein_id="CAC12904.1"
/db_xref="GI:10799537"
/translation="MSFPLRIGLSNFDGSCSCQGEAVNPYCAVLVKEVSESGNM
YIQKPPMPQWDSFTFAHINKRVMQIIYVKGKNVDLISITVELYSLAERCPKNNKG
YIQLLEKPOGMLMNAFYLEMSDITKDMNEFFGFEALHQRGGAQKQVHHVYKCH
EFTATFFPQPTFSCVCFEYVGLNKKQYQCQNAAIHKKCIDKVIKACTQGSAINRSE
TMFHKRFKIDPFRKFKVYKSPSTFEGHGTLLWGLARGLKCDACGMNVVHRCOTKE
VANLCGINOKLMARALAMISTQK"
complement(757. .868)
/gene="PRKCQ"
/note="match: GSS: Em:AQ310045"
794. .1218
/note="MER66A repeat: matches 1. .478 of consensus"
1911. .1943
/note="MIR repeat: matches 81. .115 of consensus"
2191. .2500
/note="AluX repeat: matches 1. .303 of consensus"
2661. .2701
/note="L2 repeat: matches 2651. .2691 of consensus"
2863. .3158
/note="AluSx repeat: matches 1. .298 of consensus"
3768. .3555
/note="AluJo repeat: matches 1. .287 of consensus"
3672. .3990
/note="match: STS: Em:HSJ83C6"
4921. .5255
/note="AluJo repeat: matches 3. .311 of consensus"
6886. .7058
/note="MIR repeat: matches 28. .205 of consensus"
7059. .7430
/note="match: GSS: Em:AQ125478"
7626. .7945
/note="MER33 repeat: matches 4. .324 of consensus"
8200. .8301
/note="MIR repeat: matches 20. .123 of consensus"
8362. .9708
/note="L2 repeat: matches 566. .2144 of consensus"
9826. .10196
/note="L2 repeat: matches 2315. .2745 of consensus"
10243. .10408
/note="L1P3 repeat: matches 5959. .6127 of consensus"
10409. .10716
/note="AluSx repeat: matches 1. .307 of consensus"
11141. .11319
/note="AluSg/x repeat: matches 131. .309 of consensus"
14842. .14961
/note="20 copies 6 mer atatat 67% conserved"
14873. .14968
/note="8 copies 12 mer 67% conserved"

```

```

misc_feature      15109..15456
/note="match: GSS: Em:AQ131538"
misc_feature      15123..15920
/note="match: GSS: Em:AQ097533"
repeat_region     16015..16324
/note="AluSq repeat: matches 1..308 of consensus"
repeat_region     16510..16637
/note="L2 repeat: matches 2579..2701 of consensus"
repeat_region     18266..18551
/note="Aluo repeat: matches 1..286 of consensus"
repeat_region     19367..19666
/note="AluY repeat: matches 1..301 of consensus"
repeat_region     19870..20145
/note="46 copies 6 mer ttcttt 57% conserved"
repeat_region     19871..20122
/note="7 copies 36 mer 67% conserved"
repeat_region     19872..20147
/note="23 copies 12 mer 60% conserved"
repeat_region     19890..20141
/note="3 copies 84 mer 74% conserved"
repeat_region     19897..20141
/note="5 copies 49 mer 66% conserved"
repeat_region     19985..20059
/note="15 copies 5 mer ttcttt 68% conserved"
repeat_region     20148..20418
/note="AluJb repeat: matches 3..273 of consensus"
repeat_region     20685..20902
/note="LINC5 repeat: matches 7726..7931 of consensus"
repeat_region     20910..20960
/note="U5 repeat: matches 6..55 of consensus"
repeat_region     20961..21132
/note="LINC5 repeat: matches 7566..7727 of consensus"
misc_feature      21640..22146
/note="match: GSS: Em:AQ541207"
misc_feature      complement(22776..23021)
/note="PRKCQ"
repeat_region     22859..23008
/note="MIR repeat: matches 31..191 of consensus"
repeat_region     23514..23599
/note="43 copies 2 mer aa 65% conserved"
repeat_region     23953..24183
/note="AluSq repeat: matches 1..231 of consensus"
repeat_region     24487..24558
/note="2 copies 36 mer 94% conserved"
repeat_region     25934..26379
/note="LOR1b repeat: matches 1..461 of consensus"
misc_feature      26303..26701
/note="match: GSS: Em:AQ053154"
misc_feature      27276..27777
/note="match: GSS: Em:AQ245807"
repeat_region     28312..28514
/note="L2 repeat: matches 2437..2665 of consensus"
repeat_region     32062..32097
/note="18 copies 2 mer tt 97% conserved"
repeat_region     32445..32674
/note="L1ME1 repeat: matches 5904..6168 of consensus"
repeat_region     33017..33320
/note="AluX repeat: matches 3..310 of consensus"
repeat_region     33460..33555
/note="48 copies 2 mer at 66% conserved"
repeat_region     33474..33545
/note="6 copies 12 mer 77% conserved"
repeat_region     33475..33558
/note="14 copies 6 mer tatgta 71% conserved"
repeat_region     34031..34324
/note="MER49 repeat: matches 449..738 of consensus"
misc_feature      complement(34143..34519)
/note="PRKCQ"

```

```

Query Match      6.7%; Score 24; DB 9; Length 132741;
Best Local Similarity 100.0%; Pred. No. 0.024;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 326 ggaggaggaggaggaggaggag 349  
 |||||  
 Db 19980 GGAGGAGGAGGAGGAGGAAG 19957

RESULT 44  
 AL163512  
 LOCUS  
 DEFINITION Mouse DNA sequence from clone RP21-468E12 on chromosome X, complete sequence.  
 ACCESSION AL163512  
 VERSION AL163512.35 GI:17017692  
 KEYWORDS HTG.  
 SOURCE house mouse.  
 ORGANISM Mus musculus

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

Submitted (17-NOV-2001) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: [clonerequest@sanger.ac.uk](mailto:clonerequest@sanger.ac.uk)  
 On Nov 20, 2001 this sequence version replaced gi:15808147.  
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone. Note that the only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; SW, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/projects/c\\_elegans/wormpep](http://www.sanger.ac.uk/projects/c_elegans/wormpep)

This sequence is the entire insert of clone RP21-468E12. Location/Qualifiers  
 1..134334  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /chromosome="X"  
 /clone="RP21-468E12"  
 /clone\_lib="RPCI-21"  
 15023  
 /note="Random repeat. Forced join. Gap size estimated to be approximately 280bp by restriction digest data."  
 114512  
 /note="Tandem repeat. Forced join. Gap size estimated to be approximately 90bp by restriction digest data."  
 /note="Single clone region. Assembly confirmed by restriction digest data."  
 complement(134064..134068)  
 /note="Single clone region. Assembly confirmed by restriction digest data."  
 complement(134070..134075)  
 /note="Single clone region. Assembly confirmed by restriction digest data."  
 complement(134077..134087)  
 /note="Single clone region. Assembly confirmed by restriction digest data."

## FEATURES

## source

1 4285: contig of 4285 bp in length  
 4286 4385: gap of 100 bp  
 4386 67924: contig of 63539 bp in length  
 67925 68024: gap of 100 bp  
 68025 71010: contig of 2986 bp in length  
 71011 71110: gap of 100 bp  
 71111 89837: contig of 18727 bp in length  
 89838 89937: gap of 100 bp  
 89938 113010: contig of 23073 bp in length  
 113011 113110: gap of 100 bp  
 11311 137072: contig of 23962 bp in length.  
 Location/Qualifiers  
 1..137072  
 /organism="Homo sapiens"

## FEATURES

## source

BASE COUNT 41709 a 26034 c 25611 g 40980 t  
 ORIGIN

Query Match 6.7%; Score 24; DB 10; Length 134334;  
 Best Local Similarity 100.0%; Pred. No. 0.024;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 326 ggaggaggaggaggaggaggag 349  
 |||||  
 Db 114497 GGAGGAGGAGGAGGAGGAAG 114520

## RESULT 45

## AL353893

## LOCUS

DEFINITION Homo sapiens chromosome 1 clone RP5-1108E5, \*\*\* SEQUENCING IN PROGRESS \*\*\*; 6 unordered pieces.

## ACCESSION

## AL353893

## VERSION

## AL353893.2

## KEYWORDS

## HTG; HTGS\_PHASE1; HTGS-CANCELLED.

## SOURCE

## human.

## ORGANISM

## Homo sapiens

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: [humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk)  
 On Jul 15, 2000 this sequence version replaced gi:8052073.  
 requests: [clonerequest@sanger.ac.uk](mailto:clonerequest@sanger.ac.uk)  
 On August 15, 2000 this sequence version replaced gi:8052073.  
 ----- Genome Center  
 Center: Sanger Centre  
 Center code: SC  
 Web site: <http://www.sanger.ac.uk>  
 Contact: [humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk)  
 ----- Project Information  
 Center project name: dj1108E5  
 ----- Summary Statistics  
 Assembly program: XGAP4; version 4.5  
 Sequencing vector: plasmid; L08752; 100% of reads  
 Chemistry: dye-terminator Big Dye; 100% of reads  
 Consensus quality: 134274 bases at least Q40  
 Consensus quality: 135563 bases at least Q30  
 Consensus quality: 136185 bases at least Q20  
 Insert size: 136572; sum-of-contigs  
 Quality coverage: 137237; 5.1% error; agarose-fp  
 coverage: 4.50x in Q20 bases; agarose-fp

\* NOTE: This is a 'working draft' sequence. It currently consists of 6 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 4285: contig of 4285 bp in length  
 4286 4385: gap of 100 bp  
 4386 67924: contig of 63539 bp in length  
 67925 68024: gap of 100 bp  
 68025 71010: contig of 2986 bp in length  
 71011 71110: gap of 100 bp  
 71111 89837: contig of 18727 bp in length  
 89838 89937: gap of 100 bp  
 89938 113010: contig of 23073 bp in length  
 113011 113110: gap of 100 bp  
 11311 137072: contig of 23962 bp in length.  
 Location/Qualifiers  
 1..137072  
 /organism="Homo sapiens"





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 06:06:55 ; Search time 521.76 Seconds  
(without alignments)  
1174.752 Million cell updates/sec

Title: US-09-846-456-2  
Perfect score: 357  
Sequence: 1 tggaggtctcagctgaggg.....gagggaaggagctgtgttg 357

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 1736436 seqs, 858457221 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N\_Geneseq\_032802.\*  
1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	159	44.5	10545	21 AAC69132	Human ABC1 gene ex
2	159	44.5	183999	22 AAF92831	Human ABC1 genomic
3	77	21.6	10442	22 AAF24680	Nucleotide sequenc
4	77	21.6	10442	22 AAF24702	Nucleotide sequenc
5	77	21.6	10474	22 AAF24685	Nucleotide sequenc
6	77	21.6	10474	22 AAF24686	Nucleotide sequenc
7	77	21.6	10474	22 AAF24707	Nucleotide sequenc
8	77	21.6	10474	22 AAF24708	Nucleotide sequenc
9	60	16.8	446	22 AAS04035	Partial human ABC1

10	60	16.8	7086	22 ABA09200	Human ABCA1 homolo
11	60	16.8	7086	22 AAK52667	Human polynucleoti
12	60	16.8	7260	22 AAD21326	Human ATP binding
13	60	16.8	7260	22 AAI70315	Human ATP binding
14	60	16.8	7281	22 AAK51683	Human polynucleoti
15	60	16.8	7857	21 AAC69388	Human ABC1 cholest
16	60	16.8	7860	22 AAF83826	Human ABC1 nucleot
17	60	16.8	7860	22 AAF92835	Human ABC1 cDNA.
18	60	16.8	7861	21 AAC69387	Human ABC1 cholest
19	60	16.8	7864	21 AAC69120	Human ABC1 cholest
20	60	16.8	7864	21 AAC69385	Human ABC1 cholest
21	60	16.8	7864	21 AAC69386	Human ABC1 cholest
22	60	16.8	7864	21 AAC69389	Human ABC1 cholest
23	60	16.8	9741	22 AAS06120	Human ABC1 DNA seq
24	60	16.8	9854	22 AAS06121	Human ABC1 DNA seq
25	51	14.3	736	22 AAH07432	Human cDNA clone (
26	51	14.3	1556	22 AAH18606	Human cDNA sequenc
27	35	9.8	37	22 AAF93084	ABCI polymorphism
28	35	9.8	38	22 AAF93082	ABCI polymorphism
29	24	6.7	12425	22 AAF26495	Human low density
30	22	6.2	30	21 AAZ39175	Human BMP-4 5' ups
31	22	6.2	1077	22 AAC83229	DNA sequence from
32	22	6.2	1456	19 AAV32783	Human BMP-4 promot
33	22	6.2	1751	13 AAO32853	BMP4. Rattus ratt
34	22	6.2	1751	19 AAV01680	Bone morphogenetic
35	22	6.2	1788	14 AAO53144	Sequence encoding
36	22	6.2	1788	15 AAO72710	Prepro human CBMP2
37	22	6.2	1788	17 AAT02601	Human CBMP2(b) CDN
38	22	6.2	1788	19 AAV15207	Human osteogenic p
39	22	6.2	1788	20 AAZ27580	Human prepro CBMP2
40	22	6.2	1788	20 AAX00232	Human osteogenic p
41	22	6.2	1832	23 AAS79379	DNA encoding novel
42	22	6.2	1944	18 AAT78942	Human bone morphog
43	22	6.2	1954	9 AAN80634	Human Bone Morphog
44	22	6.2	1954	12 AAO14037	Human BMP-2B in la
45	22	6.2	1954	13 AAO31870	Human BMP-2B sequ

ALIGNMENTS

RESULT 1  
AAC69132  
ID AAC69132 standard; DNA; 10545 BP.  
XX  
AC AAC69132;  
XX  
AC  
XX  
DT 29-JAN-2001 (first entry)  
XX  
DE Human ABC1 gene exon 1 (promoter).

XX KW Human ABC1 cholesterol transporter; chromosome 9q31; promoter;  
XX KW ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;  
XX KW Tangier disease; TB; familial HDL deficiency; FHA; polymorphism;  
XX KW cardiovascular disease; coronary artery disease; coronary restenosis;  
XX KW cerebrovascular disease; peripheral vascular disease;  
XX KW Alzheimer's disease; Niemann-Pick disease; Huntington's disease;  
XX KW X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;  
XX KW prognosis; prophylaxis; drug screening; transgenic animal; ss.  
XX OS Homo sapiens.

XX PN WO200055318-A2.  
XX PD 21-SEP-2000.  
XX PF 15-MAR-2000; 2000WO-IB00532.  
XX PR 15-MAR-1999; 99US-0124702.  
XX PR 08-JUN-1999; 99US-0138048.  
XX PR 17-JUN-1999; 99US-0139600.  
XX PR 01-SEP-1999; 99US-0151977.

PA (UYBR-) UNIV BRITISH COLUMBIA.  
XX (XENO-) XENON BIORESEARCH INC.  
PI Hayden MR, Wilson AR, Pimstone SN;  
XX WPI; 2000-587528/55.  
XX New ABC1 polypeptide is useful for treating diseases associated with  
PT ABC1 biological activity, e.g. Alzheimer's disease, Huntington's  
PT disease and cancer -  
XX  
XX  
PS Claim 50; Fig 12; 229pp; English.  
XX  
XX The invention relates to the human ABC1 cholesterol transporter protein  
CC (B3082) and to nucleic acid sequences (C69120) which encode it. ABC1 is  
CC a member of the ATP-binding cassette (ABC transporter) superfamily of  
CC proteins, and plays a crucial role in cholesterol transport, particularly  
CC intracellular cholesterol trafficking in monocytes and fibroblasts, being  
CC involved in cholesterol efflux from the cell. The gene encoding ABC1 is  
CC located on chromosome 9q31, and mutations in this gene are associated  
CC with two genetic HDL (high density lipoprotein) deficiency disorders,  
CC Tangier disease (TD) and familial HDL deficiency (FHA). These diseases  
CC are distinguishable in that TD is an autosomal recessive disorder, while  
CC FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good  
CC cholesterol") in the blood correlate with a high risk of cardiovascular  
CC disease, particularly coronary artery disease, but also cerebrovascular  
CC disease, coronary restenosis, and peripheral vascular disease.  
CC Conversely, a high level of HDL has protective effects against  
CC cardiovascular disease. The invention provides genetic constructs and  
CC transgenic cells and non-human animals comprising human ABC1 nucleic  
CC acids, and methods of gene therapy for the treatment or prevention of  
CC cardiovascular disease comprising the administration of an expression  
CC vector encoding ABC1 or an active fragment thereof. The invention also  
CC encompasses compounds which mimic ABC1 activity, compounds which  
CC stimulate ABC1 expression and methods of screening for such compounds.  
CC It further relates to methods for determining whether a patient has an  
CC increased risk for cardiovascular disease due to polymorphisms in the  
CC ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat  
CC or prevent cardiovascular disease, especially coronary artery disease,  
CC cerebrovascular disease, coronary restenosis or peripheral vascular  
CC disease. They may also be used in the treatment of diseases associated  
CC with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick  
CC disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.  
CC The invention specifically excludes proteins with the exact amino acid  
CC sequences of GenBank accession No: CAA10005.1 and X75926, and the nucleic  
CC acid with the exact sequence as GenBank Accession No: AJ012376.1. The  
CC present sequence represents the human ABC1 gene promoter region (exon 1).  
XX  
XX Sequence 10545 BP; 2647 A; 2225 C; 2411 G; 3256 T; 6 other;

Query Match 44.5%; Score 159; DB 21; Length 10545;  
Best Local Similarity 100.0%; Pred. No. 2.3e-68;  
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 199 ttttgccctcagctgaggttgcctgctggaagaacccacttcagagaagacaaca 258  
DB 8240 ttttgccctcagctgaggttgcctgctggaagaacccacttcagagaagacaaca 8299  
QY 259 gtaagcttgggttttccagcagcggggttctctcatttttctgtgtgtgtgtgt 318  
DB 8300 gtaagcttgggttttccagcagcggggttctctcatttttctgtgtgtgtgtgt 8359  
QY 319 ggggattggaggaggaggaggaggaggaggaggaggaggaggaggaggaggagg 357  
DB 8360 ggggattggaggaggaggaggaggaggaggaggaggaggaggaggaggaggagg 8398

RESULT 2  
ID AAF92831  
XX AAF92831 standard; DNA; 183999 BP.  
AC AAF92831;

XX 17-MAY-2001 (first entry)  
XX Human ABC1 genomic DNA.  
XX High density lipoprotein-cholesterol; HDL-C; cardiovascular; ABC1; ds.  
XX Homo sapiens.  
XX WO200115676-A2.  
XX 08-MAR-2001.  
XX 01-SEP-2000; 2000WO-IB01492.  
XX 01-SEP-1999; 99US-0151977.  
XX 15-MAR-2000; 2000US-0526193.  
XX 23-JUN-2000; 2000US-0213958.  
XX (UYBR-) UNIV BRITISH COLUMBIA.  
XX (XENO-) XENON GENETICS INC.  
XX Hayden MR, Brooks-Wilson AR, Pimstone SN, Clee SM;  
XX WPI; 2001-244356/25.  
XX Treating a lower than normal high density lipoprotein-cholesterol  
XX (HDL-C) level, a higher than normal triglyceride level, or a  
XX cardiovascular disease, by administering a compound that modulates LXR-  
XX or RXR-mediated transcriptional activity -  
XX Claim 8; Fig 1; 317pp; English.  
XX The present invention relates to a method for treating a patient  
XX diagnosed as having a lower than normal high density  
XX lipoprotein-cholesterol (HDL-C) level, a higher than normal  
XX triglyceride level, or a cardiovascular disease, involving  
XX administering a compound that modulates LXR- or RXR-mediated  
XX transcriptional activity or ABC1 expression or activity.  
XX The LXR gene product may be used in an assay to identify  
XX compounds useful for the treatment of a disease or condition selected a  
XX lower than normal HDL cholesterol level, a higher than normal  
XX triglyceride level, and a cardiovascular disease.  
XX Sequence 183999 BP; 49549 A; 37944 C; 41170 G; 54950 T; 386 other;

Query Match 44.5%; Score 159; DB 22; Length 183999;  
Best Local Similarity 100.0%; Pred. No. 2.3e-68;  
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 199 ttttgccctcagctgaggttgcctgctggaagaacccacttcagagaagacaaca 258  
DB 53328 ttttgccctcagctgaggttgcctgctggaagaacccacttcagagaagacaaca 53387  
QY 259 gtaagcttgggttttccagcagcggggttctctcatttttctgtgtgtgtgtgt 318  
DB 53388 gtaagcttgggttttccagcagcggggttctctcatttttctgtgtgtgtgtgt 53447  
QY 319 ggggattggaggaggaggaggaggaggaggaggaggaggaggaggaggaggagg 357  
DB 53448 ggggattggaggaggaggaggaggaggaggaggaggaggaggaggaggaggagg 53486

RESULT 3  
ID AAF24680  
XX AAF24680 standard; DNA; 10442 BP.  
XX AAF24680;  
XX 20-APR-2001 (first entry)  
XX Nucleotide sequence of a human ABC1 polypeptide.



XX Human; adenosine triphosphate binding cassette protein 1; ABC1;  
KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;  
KW chromosome 9q22-q31; heart disease; hypercholesterolemia;  
KW atherosclerosis; cholesterol transport; ss.  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
XX CDS 291..7076  
XX /\*tag= a  
XX /product= "ABC1 polypeptide"  
XX WO200078972-A2.  
XX  
XX 28-DEC-2000.  
XX  
XX 16-JUN-2000; 2000WO-US16765.  
XX  
XX 18-JUN-1999; 99US-0140264.  
XX 14-SEP-1999; 99US-0153872.  
XX 19-NOV-1999; 99US-0166573.  
XX (CVTH-) CV THERAPEUTICS INC.  
XX Lawn RM, Wade D, Garvin M;  
XX WPI; 2001-137812/14.  
XX Adenosine triphosphate (ATP) binding cassette (ABC) polynucleotide,  
XX useful for the development of agents for the treatment of heart disease  
XX and other disorders associated with hypercholesterolemia and  
XX atherosclerosis -  
XX  
XX Disclosure; Page 122-128; 215pp; English.  
XX  
XX The present sequence encodes a human adenosine triphosphate (ATP)  
XX binding cassette protein (ABC) 1 polypeptide. ABC1 resides in cell  
XX membranes and utilises ATP hydrolysis to transport a wide variety of  
XX substrates across the plasma membrane. ABC1 is a pivotal protein in  
XX the apolipoprotein-mediated mobilisation of intracellular cholesterol  
XX stores. ABC1 is defective in Tangier disease, a genetic disorder  
XX characterised by abnormal HDL-cholesterol metabolism. The ABC1 gene is  
XX localised to chromosome 9q22-q31. The ABC1 genes and proteins are  
XX useful for developing pharmaceutical agents for the treatment of heart  
XX disease and other disorders associated with hypercholesterolemia and  
XX atherosclerosis. The genes are useful for developing screening assays to  
XX screen for compounds that regulate the expression of genes associated  
XX with cholesterol transport. The genes and proteins are also useful for  
XX are also useful as diagnostic indicators of cardiovascular disease and  
XX other disorders associated with hypercholesterolemia.  
XX  
XX Sequence 10442 BP; 2898 A; 2297 C; 2408 G; 2835 T; 4 other;  
SQ  
Query Match 21.6%; Score 77; DB 22; Length 10442;  
Best Local Similarity 99.2%; Pred. No. 4.8e-28;  
Matches 127; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 131 agctctggccgctgcttccagggctcccgagccacacgctggcgctgctgagggga 190  
Db 229 agctctggccgctgcttccagggctcccgagccacacgctggcgctgctgagggga 288  
QY 191 acatggcatgttgctcagctaggttctgctggaagaacctcactttcagaagaa 250  
Db 289 acatggctgttgctcagctaggttctgctggaagaacctcactttcagaagaa 348  
QY 251 gacaaaca 258  
Db 349 gacaaaca 356  
RESULT 4

AAF24702

ID AAF24702 standard; DNA; 10442 BP.

XX AAF24702;

XX 20-APR-2001 (first entry)

XX Nucleotide sequence of a human ABC1 polypeptide.

XX Human; adenosine triphosphate binding cassette protein 1; ABC1;  
KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;  
KW chromosome 9q22-q31; heart disease; hypercholesterolemia;  
KW atherosclerosis; cholesterol transport; ss.  
XX  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
XX CDS 291..7076  
XX /\*tag= a  
XX /product= "ABC1 polypeptide"  
XX WO200078971-A2.  
XX  
XX 28-DEC-2000.  
XX  
XX 16-JUN-2000; 2000WO-US16591.  
XX  
XX 18-JUN-1999; 99US-0140264.  
XX 14-SEP-1999; 99US-0153872.  
XX 19-NOV-1999; 99US-0166573.  
XX (CVTH-) CV THERAPEUTICS INC.  
XX (UNIW) UNIV WASHINGTON.  
XX Lawn RM, Wade D, Oram JF, Garvin M;  
XX WPI; 2001-137811/14.  
XX P-PSDB; AAB31365.  
XX Adenosine triphosphate (ATP) binding cassette protein (ABC) 1  
XX polynucleotides and polypeptides, useful for treatment of heart disease  
XX and other disorders associated with hypercholesterolemia and  
XX atherosclerosis -  
XX  
XX Claim 3; Page 117-123; 211pp; English.  
XX  
XX The present sequence encodes a human adenosine triphosphate (ATP)  
XX binding cassette protein (ABC) 1 polypeptide. ABC1 resides in cell  
XX membranes and utilises ATP hydrolysis to transport a wide variety of  
XX substrates across the plasma membrane. ABC1 is a pivotal protein in  
XX the apolipoprotein-mediated mobilisation of intracellular cholesterol  
XX stores. ABC1 is defective in Tangier disease, a genetic disorder  
XX characterised by abnormal HDL-cholesterol metabolism. The ABC1 gene is  
XX localised to chromosome 9q22-q31. The ABC1 genes and proteins are  
XX useful for developing pharmaceutical agents for the treatment of heart  
XX disease and other disorders associated with hypercholesterolemia and  
XX atherosclerosis. The genes are useful for developing screening assays to  
XX screen for compounds that regulate the expression of genes associated  
XX with cholesterol transport. The genes and proteins are also useful for  
XX with cholesterol transport. The genes and proteins are also useful for  
XX are also useful as diagnostic indicators of cardiovascular disease and  
XX other disorders associated with hypercholesterolemia.  
XX  
XX Sequence 10442 BP; 2898 A; 2297 C; 2408 G; 2835 T; 4 other;  
SQ

Query Match 21.6%; Score 77; DB 22; Length 10442;

Best Local Similarity 99.2%; Pred. No. 4.8e-28;

Matches 127; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 131 agctctggccgctgcttccagggctcccgagccacacgctggcgctgctgagggga 190  
Db 229 agctctggccgctgcttccagggctcccgagccacacgctggcgctgctgagggga 288

QY 191 acatggcatgttgccctcagctgaggtgtgctgtggaagaacacctcacttccagaaga 250  
 Db 289 acatggctgtgtgctcagctgaggtgtgctgtggaagaacacctcacttccagaaga 348  
 QY 251 gacaaaca 258  
 Db 349 gacaaaca 356  
 RESULT 5  
 ID AAF24685  
 XX AAF24685 standard; DNA; 10474 BP.  
 AC AAF24685;  
 XX  
 DT 20-APR-2001 (first entry)  
 XX  
 DE Nucleotide sequence of ABC1 polypeptide from Tangier disease patient.  
 XX  
 DE Human; adenosine triphosphate binding cassette protein 1; ABC1;  
 KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;  
 KW chromosome 9q22-9q31; heart disease; hypercholesterolemia;  
 KW atherosclerosis; cholesterol transport; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key  
 CDS Location/Qualifiers  
 FT 323..7108  
 FT /\*tag= a  
 FT /product= "defective ABC1 polypeptide"  
 XX  
 XX WO200078972-A2.  
 XX  
 XX 28-DEC-2000.  
 XX  
 PF 16-JUN-2000; 2000WO-US16765.  
 XX  
 PR 18-JUN-1999; 99US-0140264.  
 PR 14-SEP-1999; 99US-0153872.  
 PR 19-NOV-1999; 99US-0166573.  
 XX  
 XX (CVTH-) CV THERAPEUTICS INC.  
 PA  
 PI Lawn RM, Wade D, Garvin M;  
 XX  
 DR WPI; 2001-137812/14.  
 XX  
 PT Adenosine triphosphate (ATP) binding cassette (ABC) polynucleotide,  
 PT useful for the development of agents for the treatment of heart disease  
 PT and other disorders associated with hypercholesterolemia and  
 PT atherosclerosis -  
 XX  
 PS Disclosure; Page 148-154; 215pp; English.  
 XX  
 CC The present sequence encodes a human adenosine triphosphate (ATP)  
 CC binding cassette protein (ABC) 1 polypeptide, and is isolated from  
 CC a Tangier disease patient. ABC1 resides in cell membranes and utilises  
 CC ATP hydrolysis to transport a wide variety of substrates across the  
 CC plasma membrane. ABC1 is a pivotal protein in the apolipoprotein-mediated  
 CC mobilisation of intracellular cholesterol stores. ABC1 is defective in  
 CC Tangier disease, a genetic disorder characterised by abnormal  
 CC HDL-cholesterol metabolism. The ABC1 gene is localised to chromosome  
 CC 9q22-9q31. The ABC1 genes and proteins are useful for developing  
 CC pharmaceutical agents for the treatment of heart disease and other  
 CC disorders associated with hypercholesterolemia and atherosclerosis. The  
 CC genes are useful for developing screening assays to screen for compounds  
 CC that regulate the expression of genes associated with cholesterol  
 CC transport. The genes and proteins are also useful for are also useful  
 CC as diagnostic indicators of cardiovascular disease and other disorders  
 CC associated with hypercholesterolemia.  
 XX  
 SQ Sequence 10474 BP; 2906 A; 2305 C; 2416 G; 2843 T; 4 other;

Query Match 21.6%; Score 77; DB 22; Length 10474;  
 Best Local Similarity 99.2%; Pred. No. 4.8e-28;  
 Matches 127; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 131 agctctggccgctgcttccagggctccgagccacacgctggcggtgctgagggga 190  
 Db 261 agctctggccgctgcttccagggctccgagccacacgctggcggtgctgagggga 320  
 QY 191 acatggcatgttgccctcagctgaggtgtgctgtggaagaacacctcacttccagaaga 250  
 Db 321 acatggcatgttgccctcagctgaggtgtgctgtggaagaacacctcacttccagaaga 380  
 QY 251 gacaaaca 258  
 Db 381 gacaaaca 388  
 RESULT 6  
 AAF24686  
 ID AAF24686 standard; DNA; 10474 BP.  
 XX  
 AC AAF24686;  
 XX  
 DT 20-APR-2001 (first entry)  
 XX  
 DE Nucleotide sequence of ABC1 polypeptide from Tangier disease patient.  
 XX  
 DE Human; adenosine triphosphate binding cassette protein 1; ABC1;  
 KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;  
 KW chromosome 9q22-9q31; heart disease; hypercholesterolemia;  
 KW atherosclerosis; cholesterol transport; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key  
 CDS Location/Qualifiers  
 FT 323..7108  
 FT /\*tag= a  
 FT /product= "defective ABC1 polypeptide"  
 XX  
 XX WO200078972-A2.  
 XX  
 XX 28-DEC-2000.  
 XX  
 PF 16-JUN-2000; 2000WO-US16765.  
 XX  
 PR 18-JUN-1999; 99US-0140264.  
 PR 14-SEP-1999; 99US-0153872.  
 PR 19-NOV-1999; 99US-0166573.  
 XX  
 XX (CVTH-) CV THERAPEUTICS INC.  
 PA  
 PI Lawn RM, Wade D, Garvin M;  
 XX  
 DR WPI; 2001-137812/14.  
 XX  
 PT Adenosine triphosphate (ATP) binding cassette (ABC) polynucleotide,  
 PT useful for the development of agents for the treatment of heart disease  
 PT and other disorders associated with hypercholesterolemia and  
 PT atherosclerosis -  
 XX  
 PS Disclosure; Page 170-176; 215pp; English.  
 XX  
 CC The present sequence encodes a human adenosine triphosphate (ATP)  
 CC binding cassette protein (ABC) 1 polypeptide, and is isolated from  
 CC a Tangier disease patient. ABC1 resides in cell membranes and utilises  
 CC ATP hydrolysis to transport a wide variety of substrates across the  
 CC plasma membrane. ABC1 is a pivotal protein in the apolipoprotein-mediated  
 CC mobilisation of intracellular cholesterol stores. ABC1 is defective in  
 CC Tangier disease, a genetic disorder characterised by abnormal  
 CC HDL-cholesterol metabolism. The ABC1 gene is localised to chromosome  
 CC 9q22-9q31. The ABC1 genes and proteins are useful for developing  
 CC pharmaceutical agents for the treatment of heart disease and other

CC disorders associated with hypercholesterolemia and atherosclerosis. The  
CC genes are useful for developing screening assays to screen for compounds  
CC that regulate the expression of genes associated with cholesterol  
CC transport. The genes and proteins are also useful for are also useful  
CC as diagnostic indicators of cardiovascular disease and other disorders  
CC associated with hypercholesterolemia.  
XX  
SQ Sequence 10474 BP; 2907 A; 2304 C; 2415 G; 2844 T; 4 other;  
  
Query Match 21.6%; Score 77; DB 22; Length 10474;  
Best Local Similarity 99.2%; Pred. No. 4.8e-28;  
Matches 127; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 131 agctctggcgcgtgcttccagggctcccgagccacacgctggcgctgctggctgagga 190  
Db 261 agctctggcgcgtgcttccagggctcccgagccacacgctggcgctgctggctgagga 320  
  
QY 191 acatggcatgttgctcagctgaggtgctgtgctggaagaaccttcacagaaga 250  
Db 321 acatggctgttgctcagctgaggtgctgtgctggaagaaccttcacagaaga 380  
  
QY 251 gacaaaca 258  
Db 381 gacaaaca 388  
  
RESULT 7  
AAF24707  
ID AAF24707 standard; DNA; 10474 BP.  
XX  
AC AAF24707;  
XX  
XX  
DT 20-APR-2001 (first entry)  
XX  
DE Nucleotide sequence of ABC1 polypeptide from Tangier disease patient.  
KW Human; adenosine triphosphate binding cassette protein 1; ABC1;  
KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;  
KW chromosome 9q22-9q31; heart disease; hypercholesterolemia;  
KW atherosclerosis; cholesterol transport; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
CDS 323..7108  
FT /\*tag= a  
FT /product= "defective ABC1 polypeptide"  
XX  
XX WO200078971-A2.  
XX  
XX 28-DEC-2000.  
XX  
XX 16-JUN-2000; 2000WO-US16591.  
XX  
XX 18-JUN-1999; 99US-0140264.  
PR 14-SEP-1999; 99US-0153872.  
PR 19-NOV-1999; 99US-0166573.  
XX  
XX (CVTH-) CV THERAPEUTICS INC.  
PA (UNIW ) UNIV WASHINGTON.  
XX  
PI Lawn RM, Wade D, Oram JF, Garvin M;  
XX  
XX WPI; 2001-137811/14.  
DR P-PSDB; AAB31366.  
XX  
XX Adenosine triphosphate (ATP) binding cassette protein (ABC) 1  
PT polynucleotides and polypeptides, useful for treatment of heart disease  
PT and other disorders associated with hypercholesterolemia and  
PT atherosclerosis -  
XX  
PS Claim 27; Page 144-150; 21lpp; English.

XX The present sequence encodes a human adenosine triphosphate (ATP)  
CC binding cassette protein (ABC) 1 polypeptide, and is isolated from  
CC a Tangier disease patient. ABC1 resides in cell membranes and utilises  
CC ATP hydrolysis to transport a wide variety of substrates across the  
CC plasma membrane. ABC1 is a pivotal protein in the apolipoprotein-mediated  
CC mobilisation of intracellular cholesterol stores. ABC1 is defective in  
CC Tangier disease, a genetic disorder characterised by abnormal  
CC HDL-cholesterol metabolism. The ABC1 gene is localised to chromosome  
CC 9q22-9q31. The ABC1 genes and proteins are useful for developing  
CC pharmaceutical agents for the treatment of heart disease and other  
CC disorders associated with hypercholesterolemia and atherosclerosis. The  
CC genes are useful for developing screening assays to screen for compounds  
CC that regulate the expression of genes associated with cholesterol  
CC transport. The genes and proteins are also useful for are also useful  
CC as diagnostic indicators of cardiovascular disease and other disorders  
CC associated with hypercholesterolemia.  
XX  
SQ Sequence 10474 BP; 2906 A; 2305 C; 2416 G; 2843 T; 4 other;  
  
Query Match 21.6%; Score 77; DB 22; Length 10474;  
Best Local Similarity 99.2%; Pred. No. 4.8e-28;  
Matches 127; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 131 agctctggcgcgtgcttccagggctcccgagccacacgctggcgctgctggctgagga 190  
Db 261 agctctggcgcgtgcttccagggctcccgagccacacgctggcgctgctggctgagga 320  
  
QY 191 acatggcatgttgctcagctgaggtgctgtgctggaagaaccttcacagaaga 250  
Db 321 acatggctgttgctcagctgaggtgctgtgctggaagaaccttcacagaaga 380  
  
QY 251 gacaaaca 258  
Db 381 gacaaaca 388  
  
RESULT 8  
AAF24708  
ID AAF24708 standard; DNA; 10474 BP.  
XX  
AC AAF24708;  
XX  
XX  
DT 20-APR-2001 (first entry)  
XX  
DE Nucleotide sequence of ABC1 polypeptide from Tangier disease patient.  
KW Human; adenosine triphosphate binding cassette protein 1; ABC1;  
KW apolipoprotein-mediated mobilisation; cholesterol; Tangier disease;  
KW chromosome 9q22-9q31; heart disease; hypercholesterolemia;  
KW atherosclerosis; cholesterol transport; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
CDS 323..7108  
FT /\*tag= a  
FT /product= "defective ABC1 polypeptide"  
XX  
XX WO200078971-A2.  
XX  
XX 28-DEC-2000.  
XX  
XX 16-JUN-2000; 2000WO-US16591.  
XX  
XX 18-JUN-1999; 99US-0140264.  
PR 14-SEP-1999; 99US-0153872.  
PR 19-NOV-1999; 99US-0166573.  
XX  
XX (CVTH-) CV THERAPEUTICS INC.  
PA (UNIW ) UNIV WASHINGTON.  
XX  
XX



PD 09-AUG-2001.  
XX 05-FEB-2001; 2001WO-US03800.  
PF 03-FEB-2000; 2000US-0496914.  
XX 27-APR-2000; 2000US-0560875.  
PR (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Drmanac RT;  
PI WPI; 2001-457740/49.  
XX P-PSDB; ABB11956.  
DR Human proteins and DNA encoding sequences useful for preventing,  
PT treating or ameliorating a medical condition in a mammalian subject  
PT e.g. arthritis and cancer -  
XX Claim 1; Page 833-835; 1963pp; English.  
PS Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and  
XX sequences ABA08225-ABA09574 represent nucleic acids encoding them. The  
CC invention also relates to vectors and recombinant host cells comprising a  
CC nucleotide of the invention, methods of producing the novel polypeptides,  
CC antibodies against the polypeptides, methods of detecting the nucleotides  
CC or polypeptides in a sample, and methods of identifying compounds which  
CC bind to polypeptides of the invention. Although novel, many of the  
CC polypeptides of the invention have homology to known proteins, and hence  
CC giving an insight into their probable biological activities, and hence  
CC potential therapeutic applications. The polypeptides of the invention may  
CC have various activities, including cytokine, cell proliferation or cell  
CC differentiation activities; stem cell growth factor activity;  
CC haematopoiesis regulatory activity; tissue growth activity;  
CC immunomodulatory activity; activin- or inhibin-related activities;  
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or  
CC thrombolytic activities; receptor or ligand activities; or may be  
CC involved in oncogenesis; cancer cell proliferation or metastasis.  
CC Depending on their biological activities, polypeptides and nucleotides of  
CC the invention are useful for preventing, treating or ameliorating medical  
CC conditions, e.g., by protein or gene therapy. Such conditions include  
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell  
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),  
CC proliferative retinopathy, atherosclerosis, coronary heart disease,  
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal  
CC vascular growth. Polypeptides involved with tissue regeneration and  
CC repair (or nucleic acids encoding them) may be used to promote wound  
CC healing (e.g., of burns, incisions and ulcers), while those with  
CC immunomodulatory activities may be used in the treatment of viral,  
CC bacterial and fungal infections in addition to immune disorders.  
CC Polypeptides with growth factor activity may be used in cell cultures to  
CC promote cell growth. For example, such polypeptides may be used to  
CC manipulate stem cells in culture to give rise to neuroepithelial cells  
CC that can be used to augment or replace cells damaged by illness,  
CC autoimmune disease or accidental damage. The polypeptides and nucleotides  
CC may also be used in the diagnosis of the above conditions, and in drug  
CC screening techniques. The present sequence represents a cDNA encoding a  
CC novel human polypeptide of the invention.  
XX  
SQ Sequence 7086 BP; 1773 A; 1739 C; 1859 G; 1715 T; 0 other;

Query Match 16.8%; Score 60; DB 22; Length 7086;  
Best Local Similarity 100.0%; Pred. No. 1.1e-19;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 199 ttttgccctcagctgaggttgcgtgtggaagaacctcactttcagaagaacaaca 258  
Db 310 ttttgccctcagctgaggttgcgtgtggaagaacctcactttcagaagaacaaca 369

RESULT 11  
AAK52667  
ID AAK52667 standard; cDNA; 7086 BP.

XX AAK52667;  
AC 06-NOV-2001 (first entry)  
XX Human polynucleotide SEQ ID NO 2196.  
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorder; arthritis; inflammation; ss.  
XX Homo sapiens.  
XX WO200157190-A2.  
XX 09-AUG-2001.  
XX 05-FEB-2001; 2001WO-US04098.  
XX 03-FEB-2000; 2000US-0496914.  
PR 27-APR-2000; 2000US-0560875.  
PR 20-JUN-2000; 2000US-0598075.  
PR 19-JUL-2000; 2000US-0620325.  
PR 01-SEP-2000; 2000US-0654936.  
PR 15-SEP-2000; 2000US-0663561.  
PR 20-OCT-2000; 2000US-0893325.  
PR 30-NOV-2000; 2000US-0728422.  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;  
PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;  
DR WPI; 2001-476283/51.  
DR P-PSDB; AAM79534.  
XX Nucleic acids encoding polypeptides with cytokine-like activities,  
PT useful in diagnosis and gene therapy -  
XX Claim 1; Page 4558-4560; 6221pp; English.  
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the  
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation.  
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666  
CC (AAM80020) are omitted as the relevant pages from the sequence listing  
CC were missing at the time of publication.  
XX  
SQ Sequence 7086 BP; 1773 A; 1739 C; 1859 G; 1715 T; 0 other;

Query Match 16.8%; Score 60; DB 22; Length 7086;  
Best Local Similarity 100.0%; Pred. No. 1.1e-19;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 199 ttttgccctcagctgaggttgcgtgtggaagaacctcactttcagaagaacaaca 258  
Db 310 ttttgccctcagctgaggttgcgtgtggaagaacctcactttcagaagaacaaca 369

RESULT 12  
AAD21326  
ID AAD21326 standard; DNA; 7260 BP.



Query Match 16.8%; Score 60; DB 22; Length 7260;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-19;  
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 tttggcctcagctgaggtgctgctgtggaagaacctcacttctcagaagaacaaaca 258  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 327 tttggcctcagctgaggtgctgctgtggaagaacctcacttctcagaagaacaaaca 386

## RESULT 14

AAK51683  
 ID AAK51683 standard; cDNA; 7281 BP.

XX AC AAK51683;

XX DT 06-NOV-2001 (first entry)

XX DE Human polynucleotide SEQ ID NO 228.

XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorder; arthritis; inflammation; ss.

XX OS Homo sapiens.

XX PN WO200157190-A2.

XX PD 09-AUG-2001.

XX PF 05-FEB-2001; 2001WO-US04098.

XX PR 03-FEB-2000; 2000US-0496914.

XX PR 27-APR-2000; 2000US-0560875.

XX PR 20-JUN-2000; 2000US-0598075.

XX PR 19-JUL-2000; 2000US-0620325.

XX PR 01-SEP-2000; 2000US-0654936.

XX PR 15-SEP-2000; 2000US-0663561.

XX PR 20-OCT-2000; 2000US-0693325.

XX PR 30-NOV-2000; 2000US-0728422.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;

XX PI Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

XX PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX DR WPI; 2001-476283/51.

XX DR P-PSDB; AAK78550.

XX PT Nucleic acids encoding polypeptides with cytokine-like activities,

XX PT useful in diagnosis and gene therapy -

XX PS Claim 1; Page 1086-1096; 6221pp; English.

XX CC The invention relates to polynucleotides (AAK51456-AAK53435) and the  
 CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity relating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activin/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation.

XX CC Note: records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666  
 CC (AAM80020) are omitted as the relevant pages from the sequence listing  
 CC were missing at the time of publication.

XX SQ Sequence 7281 BP; 1831 A; 1773 C; 1915 G; 1762 T; 0 other;

Query Match 16.8%; Score 60; DB 22; Length 7281;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-19;  
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 ttgtgctcagctgaggtgctgctgtggaagaacctcacttctcagaagaacaaaca 258  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 348 ttgtgctcagctgaggtgctgctgtggaagaacctcacttctcagaagaacaaaca 407

## RESULT 15

AAK69388

ID AAK69388 standard; cDNA; 7857 BP.

XX AC AAK69388;

XX DT 29-JAN-2001 (first entry)

XX DE Human ABC1 cholesterol transporter FHA-3 mutant cDNA (delta 5752-5757).

XX KW Human ABC1 cholesterol transporter; chromosome 9q31;

KW ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;

KW Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;

KW cardiovascular disease; coronary artery disease; coronary restenosis;

KW cerebrovascular disease; peripheral vascular disease;

KW Alzheimer's disease; Niemann-Pick disease; Huntington's disease;

KW X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;

KW prognosis; prophylaxis; drug screening; transgenic animal; mutant; ss.

XX OS Homo sapiens.

XX PN WO2000055318-A2.

XX PD 21-SEP-2000.

XX PR 15-MAR-2000; 2000WO-IB00532.

XX PR 15-MAR-1999; 99US-0124702.

XX PR 08-JUN-1999; 99US-0138048.

XX PR 17-JUN-1999; 99US-0139600.

XX PR 01-SEP-1999; 99US-0151977.

XX PA (UYBR-) UNIV BRITISH COLUMBIA.

XX PA (XENO-) XENON BIORESEARCH INC.

XX PI Hayden MR, Wilson AR, Pimstone SN;

XX DR WPI; 2000-587528/55.

XX DR P-PSDB; AAB38107.

XX PT New ABC1 polypeptide is useful for treating diseases associated with

XX PT ABC1 biological activity, e.g. Alzheimer's disease, Huntington's

XX PT disease and cancer -

XX PS Examples; Page -; 229pp; English.

XX CC The invention relates to the human ABC1 cholesterol transporter protein  
 CC (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is  
 CC a member of the ATP-binding cassette (ABC transporter) superfamily of  
 CC proteins, and plays a crucial role in cholesterol transport, particularly  
 CC intracellular cholesterol trafficking in monocytes and fibroblasts, being  
 CC involved in cholesterol efflux from the cell. The gene encoding ABC1 is  
 CC located on chromosome 9q31, and mutations in this gene are associated  
 CC with two genetic HDL (high density lipoprotein) deficiency disorders,  
 CC Tangier disease (TD) and familial HDL deficiency (FHA). These diseases  
 CC are distinguishable in that TD is an autosomal recessive disorder, while  
 CC FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good  
 CC cholesterol") in the blood correlate with a high risk of cardiovascular  
 CC disease, particularly coronary artery disease, but also cerebrovascular  
 CC disease, coronary restenosis, and peripheral vascular disease.  
 CC Conversely, a high level of HDL has protective effects against  
 CC cardiovascular disease. The invention provides genetic constructs and

transgenic cells and non-human animals comprising human ABC1 nucleic acids, and methods of gene therapy for the treatment or prevention of cardiovascular disease comprising the administration of an expression vector encoding ABC1 or an active fragment thereof. The invention also encompasses compounds which mimic ABC1 activity, compounds which stimulate ABC1 expression and methods of screening for such compounds. It further relates to methods for determining whether a patient has an increased risk for cardiovascular disease due to polymorphisms in the ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat or prevent cardiovascular disease, especially coronary artery disease, cerebrovascular disease, coronary restenosis or peripheral vascular disease. They may also be used in the treatment of diseases associated with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer. The invention specifically excludes proteins with the exact amino acid sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic acid with the exact sequence as GenBank Accession No: AJ012376.1. The present sequence represents cDNA encoding a mutant human ABC1 cholesterol transporter associated with an altered cholesterol level and therefore an altered risk of cardiovascular disease.

Note: The present sequence is not shown in the specification, but is derived from the native human ABC1 cDNA shown on pages 157-160.

Sequence 7857 BP; 2011 A; 1860 C; 2008 G; 1977 T; 1 other;

Query Match 16.8%; Score 60; DB 21; Length 7857;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-19;  
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 ttttgccctcagctgaggttctgctgtggaagaaacctcactttcagaagaacaaca 258  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 81 ttttgccctcagctgaggttctgctgtggaagaaacctcactttcagaagaacaaca 140

RESULT 16  
 AAF83826  
 ID AAF83826 standard; DNA; 7860 BP.  
 XX  
 AC AAF83826;  
 XX  
 DT 06-AUG-2001 (first entry)  
 XX  
 DE Human ABC1 nucleotide sequence.  
 XX  
 KW ABC1; antilipemic; cholesterol; inhibitor; low density lipoprotein;  
 KW LDL; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 CDS 75..3341  
 FT /\*tag= a  
 FT /product= "partial ABC1 protein"  
 FT /note= "the coding sequence continues beyond nucleotide  
 3341, possibly till position 6860 as identified  
 by translating the present sequence; part of the  
 corresponding protein is missing and nucleotide  
 3341 corresponds to the last amino acid residue  
 (position 1089) as indicated in the  
 specification"

WO2001132184-A2.  
 10-MAY-2001.  
 01-NOV-2000; 2000WO-US30109.  
 01-NOV-1999; 99US-0162803.  
 30-JUN-2000; 2000US-0215564.  
 (WISC ) WISCONSIN ALUMNI RES FOUND.

PI Attie AD, Cook M, Gray-Keller MP, Hayden MR, Pimstone S;  
 PI Brooks-Wilson A;  
 XX  
 XX WPT; 2001-335779/35.  
 DR P-PSDB; AABG2691.  
 XX  
 PT New method for inhibiting cholesterol uptake in the gut comprises  
 PT administration of an inhibitor of an ABC1 protein -  
 XX  
 XX Disclosure; Page 34-36; 41pp; English.  
 XX  
 CC The invention relates to a new method for inhibiting cholesterol uptake  
 CC in the gut that comprises administration of an inhibitor of an ABC1  
 CC protein. The method is useful for: lowering levels of LDL (low density  
 CC lipoprotein) cholesterol by reducing the activity of ABC1 protein in the  
 CC intestinal cells and the abundance of the ABC1 protein in the individual  
 CC by inhibiting the activity of the protein; identifying drugs that can  
 CC lower serum cholesterol and LDL levels comprises assaying the drug to  
 CC test if it can bind to an ABC1 protein; testing LDL cholesterol lowering  
 CC agents; and for modulation of ABC1 biological activity. The present  
 CC sequence represents a human ABC1 nucleotide sequence.  
 XX  
 XX Sequence 7860 BP; 2013 A; 1861 C; 2009 G; 1977 T; 0 other;

Query Match 16.8%; Score 60; DB 22; Length 7860;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-19;  
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 ttttgccctcagctgaggttctgctgtggaagaaacctcactttcagaagaacaaca 258  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 81 ttttgccctcagctgaggttctgctgtggaagaaacctcactttcagaagaacaaca 140

RESULT 17  
 AAF92835  
 ID AAF92835 standard; DNA; 7860 BP.  
 XX  
 AC AAF92835;  
 XX  
 DT 17-MAY-2001 (first entry)  
 XX  
 DE Human ABC1 cDNA.  
 XX  
 KW High density lipoprotein-cholesterol; HDL-C; cardiovascular; ABC1; ss.  
 KW Homo sapiens.  
 OS  
 XX WO200115676-A2.  
 XX  
 PD 08-MAR-2001.  
 XX  
 PF 01-SEP-2000; 2000WO-IB01492.  
 XX  
 PR 01-SEP-1999; 99US-0151977.  
 PR 15-MAR-2000; 2000US-0526193.  
 PR 23-JUN-2000; 2000US-0213958.  
 XX  
 PA (UYBR-) UNIV BRITISH COLUMBIA.  
 PA (XENO-) XENON GENETICS INC.  
 XX  
 PI Hayden MR, Brooks-Wilson AR, Pimstone SN, Clee SM;  
 DR WPI; 2001-244356/25.  
 XX  
 XX Treating a lower than normal high density lipoprotein-cholesterol  
 PT (HDL-C) level, a higher than normal triglyceride level, or a  
 PT cardiovascular disease, by administering a compound that modulates LXR-  
 PT or RXR-mediated transcriptional activity -  
 XX  
 PS Disclosure; Fig 2; 317pp; English.  
 XX  
 CC The present invention relates to a method for treating a patient



diagnosed as having a lower than normal high density lipoprotein-cholesterol (HDL-C) level, a higher than normal triglyceride level, or a cardiovascular disease, involving administering a compound that modulates LXR- or RXR-mediated transcriptional activity or ABC1 expression or activity.

The LXR gene product may be used in an assay to identify compounds useful for the treatment of a disease or condition selected a lower than normal HDL cholesterol level, a higher than normal triglyceride level, and a cardiovascular disease.

Sequence 7860 BP; 2014 A; 1860 C; 2008 G; 1978 T; 0 other;

Query Match 16.8%; Score 60; DB 22; Length 7860;  
Best Local Similarity 100.0%; Pred. No. 1.1e-19;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 tgttgccctcagctgaggttgctgtggaagaacctcactttcagaagaagacaaaca 258  
|||||  
Db 81 tgttgccctcagctgaggttgctgtggaagaacctcactttcagaagaagacaaaca 140

RESULT 18  
AAC69387  
ID AAC69387 standard; cDNA; 7861 BP.  
AC AAC69387;  
XX  
XX  
XX 29-JAN-2001 (first entry)  
XX Human ABC1 cholesterol transporter FHA-1 mutant cDNA (delta 2151-2153).  
XX  
XX Human ABC1 cholesterol transporter; chromosome 9q31;  
XX ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;  
XX Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;  
XX cardiovascular disease; coronary artery disease; coronary restenosis;  
XX cerebrovascular disease; peripheral vascular disease;  
XX Alzheimer's disease; Niemann-Pick disease; Huntington's disease;  
XX X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;  
XX prognosis; prophyllaxis; drug screening; transgenic animal; mutant; ss.  
XX Homo sapiens.  
XX OS  
XX WO200055318-A2.  
XX PN  
XX 21-SEP-2000.  
XX PD  
XX 15-MAR-2000; 2000WO-IB00532.  
XX PF  
XX

15-MAR-2000; 2000WO-IB00532.  
15-MAR-1999; 99US-0124702.  
08-JUN-1999; 99US-0138048.  
17-JUN-1999; 99US-0139600.  
01-SEP-1999; 99US-0151977.  
(UYBR-) UNIV BRITISH COLUMBIA.  
(XENO-) XENON BIORESEARCH INC.  
Hayden MR, Wilson AR, Pimstone SN;  
WPI; 2000-587528/55.  
P-PSDB; AAB38106.  
New ABC1 polypeptide is useful for treating diseases associated with ABC1 biological activity, e.g. Alzheimer's disease, Huntington's disease and cancer -  
Examples; Page -: 229pp; English.  
PS  
PS The invention relates to the human ABC1 cholesterol transporter protein (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is a member of the ATP-binding cassette (ABC transporter) superfamily of proteins, and plays a crucial role in cholesterol transport, particularly intracellular cholesterol trafficking in monocytes and fibroblasts, being

involved in cholesterol efflux from the cell. The gene encoding ABC1 is located on chromosome 9q31, and mutations in this gene are associated with two genetic HDL (high density lipoprotein) deficiency disorders. Tangier disease (TD) and familial HDL deficiency (FHA). These diseases are distinguishable in that TD is an autosomal recessive disorder, while FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good cholesterol") in the blood correlate with a high risk of cardiovascular disease, particularly coronary artery disease, but also cerebrovascular disease, coronary restenosis, and peripheral vascular disease. Conversely, a high level of HDL has protective effects against cardiovascular disease. The invention provides genetic constructs and transgenic cells and non-human animals comprising human ABC1 nucleic acids, and methods of gene therapy for the treatment or prevention of cardiovascular disease comprising the administration of an expression vector encoding ABC1 or an active fragment thereof. The invention also encompasses compounds which mimic ABC1 activity, compounds which stimulate ABC1 expression and methods of screening for such compounds. It further relates to methods for determining whether a patient has an increased risk for cardiovascular disease due to polymorphisms in the ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat or prevent cardiovascular disease, especially coronary artery disease, cerebrovascular disease, coronary restenosis or peripheral vascular disease. They may also be used in the treatment of diseases associated with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer. The invention specifically excludes proteins with the exact amino acid sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic acid with the exact sequence as GenBank Accession No: AJ012376.1. The present sequence represents cDNA encoding a mutant human ABC1 cholesterol transporter associated with an altered cholesterol level and therefore an altered risk of cardiovascular disease.  
Note: The present sequence is not shown in the specification, but is derived from the native human ABC1 cDNA shown on pages 157-160.

Sequence 7861 BP; 2014 A; 1859 C; 2011 G; 1976 T; 1 other;

Query Match 16.8%; Score 60; DB 21; Length 7861;  
Best Local Similarity 100.0%; Pred. No. 1.1e-19;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 tgttgccctcagctgaggttgctgtggaagaacctcactttcagaagaagacaaaca 258  
|||||  
Db 81 tgttgccctcagctgaggttgctgtggaagaacctcactttcagaagaagacaaaca 140

RESULT 19  
AAC69120  
ID AAC69120 standard; cDNA; 7864 BP.  
XX  
XX AAC69120;  
XX  
XX 29-JAN-2001 (first entry)  
XX Human ABC1 cholesterol transporter cDNA.  
XX  
XX Human ABC1 cholesterol transporter; chromosome 9q31;  
XX ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;  
XX Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;  
XX cardiovascular disease; coronary artery disease; coronary restenosis;  
XX cerebrovascular disease; peripheral vascular disease;  
XX Alzheimer's disease; Niemann-Pick disease; Huntington's disease;  
XX X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;  
XX prognosis; prophyllaxis; drug screening; transgenic animal; ss.  
XX Homo sapiens.  
XX OS  
XX WO200055318-A2.  
XX PN  
XX 21-SEP-2000.  
XX PD  
XX 15-MAR-2000; 2000WO-IB00532.  
XX PF  
XX

PR 15-MAR-1999; 99US-0124702.  
 PR 08-JUN-1999; 99US-0138048.  
 PR 17-JUN-1999; 99US-0139600.  
 PR 01-SEP-1999; 99US-0151977.  
 XX  
 PA (UYBR-) UNIV BRITISH COLUMBIA.  
 PA (XENO-) XENON BIORESEARCH INC.  
 XX  
 PI Hayden MR, Wilson AR, Pimstone SN;  
 XX  
 XX WPI: 2000-587528/55.  
 DR P-PSDB; AAB38082.  
 XX  
 PT New ABC1 polypeptide is useful for treating diseases associated with  
 PT ABC1 biological activity, e.g. Alzheimer's disease, Huntington's  
 PT disease and cancer.  
 XX  
 PS Claim 13; Page 157-160; 229pp; English.  
 XX  
 CC The invention relates to the human ABC1 cholesterol transporter protein  
 CC (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is  
 CC a member of the ATP-binding cassette (ABC transporter) superfamily of  
 CC proteins, and plays a crucial role in cholesterol transport, particularly  
 CC intracellular cholesterol trafficking in monocytes and fibroblasts, being  
 CC involved in cholesterol efflux from the cell. The gene encoding ABC1 is  
 CC located on chromosome 9q31, and mutations in this gene are associated  
 CC with two genetic HDL (high density lipoprotein) deficiency disorders,  
 CC Tangier disease (TD) and familial HDL deficiency (FHA). These diseases  
 CC are distinguishable in that TD is an autosomal recessive disorder, while  
 CC FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good  
 CC cholesterol") in the blood correlate with a high risk of cardiovascular  
 CC disease, particularly coronary artery disease, but also cerebrovascular  
 CC disease, coronary restenosis, and peripheral vascular disease.  
 CC Conversely, a high level of HDL has protective effects against  
 CC cardiovascular disease. The invention provides genetic constructs and  
 CC transgenic cells and non-human animals comprising human ABC1 nucleic  
 CC acids, and methods of gene therapy for the treatment or prevention of  
 CC cardiovascular disease comprising the administration of an expression  
 CC vector encoding ABC1 or an active fragment thereof. The invention also  
 CC encompasses compounds which mimic ABC1 activity, compounds which  
 CC stimulate ABC1 expression and methods of screening for such compounds.  
 CC It further relates to methods for determining whether a patient has an  
 CC increased risk for cardiovascular disease due to polymorphisms in the  
 CC ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat  
 CC or prevent cardiovascular disease, especially coronary artery disease,  
 CC cerebrovascular disease, coronary restenosis or peripheral vascular  
 CC disease. They may also be used in the treatment of diseases associated  
 CC with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick  
 CC disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.  
 CC The invention specifically excludes proteins with the exact amino acid  
 CC sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic  
 CC acid with the exact sequence as GenBank Accession No: AJ012376.1. The  
 CC present sequence represents cDNA encoding the human ABC1 cholesterol  
 CC transporter.  
 XX  
 SO Sequence 7864 BP: 2014 A; 1860 C; 2011 G; 1978 T; 1 other;

Query Match 16.8%; Score 60; DB 21; Length 7864;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-19;  
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 199 tgttgccagctgaggttgctgtgtggaagaacctcaacttctcagaagaacaaca 258  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 81 tgttgccagctgaggttgctgtgtggaagaacctcaacttctcagaagaacaaca 140  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 20  
 AAC69385  
 ID AAC69385 standard; cDNA; 7864 BP.  
 XX  
 AC AAC69385;  
 XX

DT 29-JAN-2001 (first entry)  
 XX Human ABC1 cholesterol transporter TD-1 mutant cDNA (T4503C).  
 DE  
 XX Human ABC1 cholesterol transporter; chromosome 9q31;  
 KW ATP-binding cassette; HDL deficiency disorder; high density lipoprotein;  
 KW Tangier disease; TD; familial HDL deficiency; FHA; polymorphism;  
 KW cardiovascular disease; coronary artery disease; coronary restenosis;  
 KW cerebrovascular disease; peripheral vascular disease;  
 KW Alzheimer's disease; Niemann-Pick disease; Huntington's disease;  
 KW X-linked adrenoleukodystrophy; cancer; gene therapy; genetic diagnosis;  
 KW prognosis; prophylaxis; drug screening; transgenic animal; mutant; ss.  
 XX Homo sapiens.  
 OS  
 XX W0200055318-A2.  
 PN  
 XX 21-SEP-2000.  
 PD  
 XX 15-MAR-2000; 2000WO-IB00532.  
 PF  
 XX 15-MAR-1999; 99US-0124702.  
 PR 08-JUN-1999; 99US-0138048.  
 PR 17-JUN-1999; 99US-0139600.  
 PR 01-SEP-1999; 99US-0151977.  
 XX  
 PA (UYBR-) UNIV BRITISH COLUMBIA.  
 PA (XENO-) XENON BIORESEARCH INC.  
 XX  
 PI Hayden MR, Wilson AR, Pimstone SN;  
 XX  
 XX WPI: 2000-587528/55.  
 DR P-PSDB; AAB38104.  
 XX  
 PT New ABC1 polypeptide is useful for treating diseases associated with  
 PT ABC1 biological activity, e.g. Alzheimer's disease, Huntington's  
 PT disease and cancer.  
 XX  
 PS Examples; Page -: 229pp; English.  
 XX  
 CC The invention relates to the human ABC1 cholesterol transporter protein  
 CC (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is  
 CC a member of the ATP-binding cassette (ABC transporter) superfamily of  
 CC proteins, and plays a crucial role in cholesterol transport, particularly  
 CC intracellular cholesterol trafficking in monocytes and fibroblasts, being  
 CC involved in cholesterol efflux from the cell. The gene encoding ABC1 is  
 CC located on chromosome 9q31, and mutations in this gene are associated  
 CC with two genetic HDL (high density lipoprotein) deficiency disorders,  
 CC Tangier disease (TD) and familial HDL deficiency (FHA). These diseases  
 CC are distinguishable in that TD is an autosomal recessive disorder, while  
 CC FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good  
 CC cholesterol") in the blood correlate with a high risk of cardiovascular  
 CC disease, particularly coronary artery disease, but also cerebrovascular  
 CC disease, coronary restenosis, and peripheral vascular disease.  
 CC Conversely, a high level of HDL has protective effects against  
 CC cardiovascular disease. The invention provides genetic constructs and  
 CC transgenic cells and non-human animals comprising human ABC1 nucleic  
 CC acids, and methods of gene therapy for the treatment or prevention of  
 CC cardiovascular disease comprising the administration of an expression  
 CC vector encoding ABC1 or an active fragment thereof. The invention also  
 CC encompasses compounds which mimic ABC1 activity, compounds which  
 CC stimulate ABC1 expression and methods of screening for such compounds.  
 CC It further relates to methods for determining whether a patient has an  
 CC increased risk for cardiovascular disease due to polymorphisms in the  
 CC ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat  
 CC or prevent cardiovascular disease, especially coronary artery disease,  
 CC cerebrovascular disease, coronary restenosis or peripheral vascular  
 CC disease. They may also be used in the treatment of diseases associated  
 CC with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick  
 CC disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.  
 CC The invention specifically excludes proteins with the exact amino acid  
 CC sequences of GenBank Accession No: CAA10005.1 and X75926, and the nucleic  
 CC acid with the exact sequence as GenBank Accession No: AJ012376.1. The  
 CC present sequence represents cDNA encoding the human ABC1 cholesterol  
 CC transporter.

The invention relates to the human ABC1 cholesterol transporter protein (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is a member of the ATP-binding cassette (ABC transporter) superfamily of proteins, and plays a crucial role in cholesterol transport, particularly intracellular cholesterol trafficking in monocytes and fibroblasts, being involved in cholesterol efflux from the cell. The gene encoding ABC1 is located on chromosome 9q31, and mutations in this gene are associated with two genetic HDL (high density lipoprotein) deficiency disorders, Tangier disease (TD) and familial HDL deficiency (FHA). These diseases

XX (UYBR-) UNIV BRITISH COLUMBIA.  
 PA (XENO-) XENON BIORESEARCH INC.  
 XX  
 PI Hayden MR, Wilson AR, Pimstone SN;  
 XX  
 XX WPI: 2000-587528/55.  
 DR P-PSDB; AAB38108.  
 XX  
 PT New ABC1 polypeptide is useful for treating diseases associated with  
 PT ABC1 biological activity, e.g. Alzheimer's disease, Huntington's  
 PT disease and cancer -  
 XX  
 PS Examples; Page -: 229pp; English.  
 XX  
 CC The invention relates to the human ABC1 cholesterol transporter protein  
 CC (B38082) and to nucleic acid sequences (C69120) which encode it. ABC1 is  
 CC a member of the ATP-binding cassette (ABC transporter) superfamily of  
 CC proteins, and plays a crucial role in cholesterol transport, particularly  
 CC intracellular cholesterol trafficking in monocytes and fibroblasts, being  
 CC involved in cholesterol efflux from the cell. The gene encoding ABC1 is  
 CC located on chromosome 9q31, and mutations in this gene are associated  
 CC with two genetic HDL (high density lipoprotein) deficiency disorders,  
 CC Tangier disease (TD) and familial HDL deficiency (FHA). These diseases  
 CC are distinguishable in that TD is an autosomal recessive disorder, while  
 CC FHA is inherited as an autosomal dominant trait. Low levels of HDL ("good  
 CC cholesterol") in the blood correlate with a high risk of cardiovascular  
 CC disease, particularly coronary artery disease, but also cerebrovascular  
 CC disease, coronary restenosis, and peripheral vascular disease.  
 CC Conversely, a high level of HDL has protective effects against  
 CC cardiovascular disease. The invention provides genetic constructs and  
 CC transgenic cells and non-human animals comprising human ABC1 nucleic  
 CC acids, and methods of gene therapy for the treatment or prevention of  
 CC cardiovascular disease comprising the administration of an expression  
 CC vector encoding ABC1 or an active fragment thereof. The invention also  
 CC encompasses compounds which mimic ABC1 activity, compounds which  
 CC stimulate ABC1 expression and methods of screening for such compounds.  
 CC It further relates to methods for determining whether a patient has an  
 CC increased risk for cardiovascular disease due to polymorphisms in the  
 CC ABC1 gene. Human ABC1 proteins and nucleotides can be used to treat  
 CC or prevent cardiovascular disease, especially coronary artery disease,  
 CC cerebrovascular disease, coronary restenosis or peripheral vascular  
 CC disease. They may also be used in the treatment of diseases associated  
 CC with ABC1 biological activity, such as Alzheimer's disease, Niemann-Pick  
 CC disease, Huntington's disease, X-linked adrenoleukodystrophy and cancer.  
 CC The invention specifically excludes proteins with the exact amino acid  
 CC sequences of GenBank accession No: CAA10005.1 and X75926, and the nucleic  
 CC acid with the exact sequence as GenBank Accession No: AJ012376.1. The  
 CC present sequence represents cDNA encoding a mutant human ABC1 cholesterol  
 CC transporter associated with an altered cholesterol level and therefore an  
 CC altered risk of cardiovascular disease.  
 CC Note: The present sequence is not shown in the specification, but is  
 CC derived from the native human ABC1 cDNA shown on pages 157-160.  
 XX  
 SQ Sequence 7864 BP; 2014 A; 1859 C; 2011 G; 1979 T; 1 other;

Query Match 16.8%; Score 60; DB 21; Length 7864;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-19;  
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 199 ttttgccctcagctgaggttgcgtgtggaagaacctcactttcagaagaacaaca 258  
 Db 81 ttttgccctcagctgaggttgcgtgtggaagaacctcactttcagaagaacaaca 140  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 23  
 AAS06120  
 ID AAS06120 standard; cDNA; 9741 BP.  
 XX  
 XX AAS06120;  
 XX  
 DT 12-SEP-2001 (first entry)

XX Human ABC1 DNA sequence #1.  
 DE  
 XX  
 KW Human; ABC1 gene; atherosclerosis; reverse transport; cholesterol;  
 KW cardiovascular; neurological; Tangier disease;; LCAT deficiency;  
 KW lecithin-cholesterol acetyltransferase; malaria; diabetes; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT 185..6967  
 FT /\*tag= a  
 FT /product= "Human ABC1 protein"  
 XX  
 PN WO200130848-A2.  
 XX  
 PD 03-MAY-2001.  
 XX  
 PF 26-OCT-2000; 2000WO-EP10886.  
 XX  
 PR 26-OCT-1999; 99EP-0402668.  
 PR 01-MAR-2000; 2000US-0186260.  
 XX  
 PA (AVET ) AVENTIS PHARMA SA.  
 XX  
 PI Denefle P, Rosier-Montus M, Arnould-Requigne L, Prades C, Naudin L;  
 PI Lemoine C, Duverger N, Jaye M, Searfoss GH, Remaley A, Brewer HB;  
 PI Dean M;  
 XX  
 XX WPI: 2001-316327/33.  
 DR P-PSDB; AAU02176.  
 XX  
 XX New human ABC1 nucleic acids and polypeptides for treating  
 PT atherosclerosis, malaria and diabetes -  
 XX  
 PS Claim 1; Page 204-208; 368pp; English.  
 XX  
 CC The sequence represents the coding sequence #1 of human ABC1. The  
 CC nucleic acid sequence, primers and probes derived from the ABC1 sequence,  
 CC and polypeptides and vectors are useful for the prevention of  
 CC atherosclerosis, in a subject affected by a dysfunction in the reverse  
 CC transport of cholesterol. The polypeptide encoded by the ABC1 gene is  
 CC useful for screening for an active ingredient for the prevention or  
 CC treatment of a disease resulting from dysfunction in the reverse  
 CC transport of cholesterol. The nucleic acids and polypeptides are also  
 CC useful for treating and preventing cardiovascular and neurological  
 CC pathologies, and other diseases e.g. Tangier disease, lecithin-  
 CC cholesterol (LCAT) deficiency, malaria and diabetes.  
 XX  
 SQ Sequence 9741 BP; 2650 A; 2180 C; 2290 G; 2620 T; 1 other;

Query Match 16.8%; Score 60; DB 22; Length 9741;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-19;  
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 199 ttttgccctcagctgaggttgcgtgtggaagaacctcactttcagaagaacaaca 258  
 Db 191 ttttgccctcagctgaggttgcgtgtggaagaacctcactttcagaagaacaaca 250  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 24  
 AAS06121  
 ID AAS06121 standard; cDNA; 9854 BP.  
 XX  
 XX AAS06121;  
 XX  
 DT 12-SEP-2001 (first entry)  
 XX  
 DE Human ABC1 DNA sequence #2.  
 XX  
 KW Human; ABC1 gene; atherosclerosis; reverse transport; cholesterol;  
 KW cardiovascular; neurological; Tangier disease;; LCAT deficiency;

KW lecithin-cholesterol acetyltransferase; malaria; diabetes; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 298..7078

FT /\*tag= a

FT /product= "Human ABC1 protein"

XX WO200130848-A2.

XX 03-MAY-2001.

XX 26-OCT-2000; 2000WO-EP10886.

XX 26-OCT-1999; 99EP-0402668.

XX 01-MAR-2000; 2000US-0186260.

XX (AVET ) AVENTIS PHARMA SA.

PI Deneffe P, Rosier-Montus M, Arnould-Reguigne I, Prades C, Naudin L;

PI Lemoine C, Duverger N, Jaye M, Searfoss GH, Remaley A, Brewer HB;

PI Dean M;

DR WPI; 2001-316327/33.

DR P-PSDB; AAU02176.

XX New human ABC1 nucleic acids and polypeptides for treating

PT atherosclerosis, malaria and diabetes -

XX Claim 1; Page 209-213; 368pp; English.

PS The sequence represents the coding sequence #2 of human ABC1. The  
CC nucleic acid sequence, primers and probes derived from the ABC1 sequence,  
CC and polypeptides and vectors are useful for the prevention of  
CC atherosclerosis, in a subject affected by a dysfunction in the reverse  
CC transport of cholesterol. The polypeptide encoded by the ABC1 gene is  
CC useful for screening for an active ingredient for the prevention or  
CC treatment of a disease resulting from dysfunction in the reverse  
CC transport of cholesterol. The nucleic acids and polypeptides are also  
CC useful for treating and preventing cardiovascular and neurological  
CC pathologies, and other diseases e.g. Tangier disease, lecithin-  
CC cholesterol (LCAT) deficiency, malaria and diabetes.

XX Sequence 9854 BP; 2665 A; 2219 C; 2334 G; 2635 T; 1 other;

Query Match 16.8%; Score 60; DB 22; Length 9854;

Best Local Similarity 100.0%; Pred. No. 1.1e-19;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 199 ttttgccctcagctgaggttgctgtgtggaagaacccctcactttcagaagaagacaaca 258

|||||

Db 304 ttttgccctcagctgaggttgctgtgtggaagaacccctcactttcagaagaagacaaca 363

RESULT 25

AAH07432

ID AAH07432 standard; cDNA; 736 BP.

XX AAH07432;

XX 26-JUN-2001 (first entry)

DE Human cDNA clone (5'-primer) SEQ ID NO:4267.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX

PF

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

PR

XX

28-JUL-2000; 2000EP-0116126.

29-JUL-1999; 99JP-0248036.

27-AUG-1999; 99JP-0300253.

11-JAN-2000; 2000JP-0118776.

02-MAY-2000; 2000JP-0183767.

09-JUN-2000; 2000JP-0241899.

(HELI-) HELIX RES INST.

Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602  
PT full-length cDNAs defined in the specification, and for the detection  
PT and/or diagnosis of the abnormality of the proteins encoded by the  
PT full-length cDNAs -

XX Claim 1; SEQ ID 4267; 2537pp + CD ROM; English.

PS The present invention describes primer sets for synthesizing 5602  
CC full-length cDNAs defined in the specification. Where a primer set  
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
CC to the complementary strand of a polynucleotide which comprises one of  
CC the 5602 nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in  
CC the specification. The primer sets can be used in antisense therapy and  
CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to  
CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632  
CC represent oligonucleotides, all of which are used in the exemplification  
CC of the present invention.

XX Sequence 736 BP; 163 A; 199 C; 199 G; 170 T; 5 other;

Query Match 14.3%; Score 51; DB 22; Length 736;

Best Local Similarity 100.0%; Pred. No. 2.9e-15;

Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 208 cagctgaggttgctgtgtggaagaacctcactttcagaagaagacaaca 258

|||||

Db 329 cagctgaggttgctgtgtggaagaacctcactttcagaagaagacaaca 379

RESULT 26

AAH18606

ID AAH18606 standard; cDNA; 1556 BP.

XX AAH18606;

XX 26-JUN-2001 (first entry)

DE Human cDNA sequence SEQ ID NO:18808.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX PD 07-FEB-2001.  
 XX PF 28-JUL-2000; 2000EP-0116126.  
 XX PR 29-JUL-1999; 99JP-0248036.  
 PR 17-AUG-1999; 99JP-0300253.  
 PR 11-JAN-2000; 2000JP-0118776.  
 PR 02-MAY-2000; 2000JP-0183767.  
 PR 09-JUN-2000; 2000JP-0241899.  
 XX PR (HELI-) HELIX RES INST.  
 XX PA Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX PT WPI; 2001-318749/34.  
 XX DR  
 XX PR Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs -  
 XX PS Claim 8; SEQ ID 18808; 2537pp + CD ROM; English.  
 XX CC The present invention describes primer sets for synthesising 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC in gene therapy. The primers are useful for synthesising polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.  
 XX SQ Sequence 1556 BP; 380 A; 363 C; 399 G; 414 T; 0 other;

Query Match 14.3%; Score 51; DB 22; Length 1556;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-15;  
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 208 cagctgaggttgcctgtggaagaacctcacttcagaagaagacaaca 258  
 ||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 329 cagctgaggttgcctgtggaagaacctcacttcagaagaagacaaca 379

RESULT 27  
 AAF93084  
 ID AAF93084 standard; DNA; 37 BP.  
 XX AC AAF93084;  
 XX DT 17-MAY-2001 (first entry)  
 XX DE ABC1 polymorphism RFLP oligonucleotide #45.  
 XX KW High density lipoprotein-cholesterol; HDL-C; cardiovascular; ABC1; ds.  
 XX OS Homo sapiens.

XX WO200115676-A2.  
 XX PD 08-MAR-2001.  
 XX XX 01-SEP-2000; 2000WO-IB01492.  
 XX PF 01-SEP-1999; 99US-0151977.  
 XX PR 15-MAR-2000; 2000US-0526193.  
 XX PR 23-JUN-2000; 2000US-0213958.  
 XX XX (UYBR-) UNIV BRITISH COLUMBIA.  
 XX PA (XENO-) XENON GENETICS INC.  
 XX PI Hayden MR, Brooks-Wilson AR, Pimstone SN, Clee SM;  
 XX PT WPI; 2001-244356/25.  
 XX DR  
 XX PR Treating a lower than normal high density lipoprotein-cholesterol  
 PT (HDL-C) level, a higher than normal triglyceride level, or a  
 PT cardiovascular disease, by administering a compound that modulates LXR-  
 PT or RXR-mediated transcriptional activity -  
 XX PS Disclosure; Fig 17; 317pp; English.  
 XX CC The present invention relates to a method for treating a patient  
 CC diagnosed as having a lower than normal high density  
 CC lipoprotein-cholesterol (HDL-C) level, a higher than normal  
 CC triglyceride level, or a cardiovascular disease, involving  
 CC administering a compound that modulates LXR- or RXR-mediated  
 CC transcriptional activity or ABC1 expression or activity.  
 CC The LXR gene product may be used in an assay to identify  
 CC compounds useful for the treatment of a disease or condition selected a  
 CC lower than normal HDL cholesterol level, a higher than normal  
 CC triglyceride level, and a cardiovascular disease.  
 XX SQ Sequence 37 BP; 4 A; 17 C; 11 G; 5 T; 0 other;  
 Query Match 9.8%; Score 35; DB 22; Length 37;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-07;  
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 138 gccgcgtccttcagggtcccgagccacacgctg 172  
 ||||||||||||||||||||||||||||||||||||||||||||  
 Db 1 gccgcgtccttcagggtcccgagccacacgctg 35  
 RESULT 28  
 AAF93082  
 ID AAF93082 standard; DNA; 38 BP.  
 XX AC AAF93082;  
 XX DT 17-MAY-2001 (first entry)  
 XX XX ABC1 polymorphism RFLP oligonucleotide #43.  
 XX DE High density lipoprotein-cholesterol; HDL-C; cardiovascular; ABC1; ds.  
 XX KW Homo sapiens.  
 XX OS WO200115676-A2.  
 XX PD 08-MAR-2001.  
 XX XX 01-SEP-2000; 2000WO-IB01492.  
 XX PF 01-SEP-1999; 99US-0151977.  
 XX PR 15-MAR-2000; 2000US-0526193.  
 XX PR 23-JUN-2000; 2000US-0213958.  
 XX XX (UYBR-) UNIV BRITISH COLUMBIA.







```

FT
/number= 1
/note= "partial intron sequence"
FT

```

AA WO9823740-A1.  
PN

XX  
PD  
04-JUN-1998.

XX 20-NOV-1997: 97WQ-EP066668.

XX  
PP  
22-NOV-1996.  
96EP-0203283.

XX  
BA /AT.KH \ AKZO NOBEL. NV

XX  
PI  
Dijkema P  
Van Den Wijngaard A:

XX  
WP: 1000-322730/38

XX Promoter from human bone morphogenetic protein-4 gene - and related  
PT vectors and transformed cells, useful for screening agents for the  
PT treatment of osteoporosis

XX  
PS  
Claim 2: Fig 5a: 41pp: English.

The present sequence represents a human bone morphogenetic protein-4 promoter (BMP-4) region 1. The invention provides a method for the identification of therapeutic agents for use in the prevention and/or treatment of osteoporosis. The method involves introducing two expression vectors into a host cell. The first expression vector comprises a BMP-4 promoter region 1 or 2 (AAV32784) which is operably linked to a reporter gene, e.g. BMP-4 gene. The second expression vector comprises of a DNA encoding an estrogen receptor. The host cell is then contacted with the compounds being tested for therapeutic activity. Compounds which bind and form a complex with the estrogen receptor, to gain entry into the cell, and then induce the expression of the reporter gene, through the BMP-4 promoter region, are potential therapeutic candidates. By this method, the antitumour agent ICI164384 was shown to be an effective therapeutic agent. Unlike oestrogens currently used to treat osteoporosis, the therapeutic agents are claimed not to have an increment in the risk of developing breast or endometrial cancer.

XX sequence 1456 BP: 357 A: 361 C: 410 G: 328 T: 0 other;

Query Match	6.2%	Score 22;	DB 19;	Length 1456;
Best Local Similarity	100.0%;	Pred. No. 0.52;		
Matches	22.	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0;

Qy 326 ggaggaggggagggaagga 347  
pb 1187 ccaagggagggaggggaagga 1208

RESULT 33  
AAQ32853  
ID AAQ32853 standard: cDNA: 1751 BP.

XX  
ACXX  
DT 05-MAY-1993 (first entry)XX  
DF  
BMP4

XX Bone; morphogenetic; protein; BMP; growth; vitamin D; systemic;  
KW treatment; dimer; ss.  
KW

XX  
OS  
Rattus rattus.

XX PN W09221365-A

XX  
PD  
10-DEC-1992

XX  
DE 26-MAY-1992: 92W0-PCS04356

XX	05-JUN-1991;	91US-0709621.
XX	PR	
XX	27-MAR-1992;	92US-0856110.
XX	PA	
XX	(PROC )	PROCTER & GAMBLE CO.
XX	Stone RL;	
XX	PI	
XX	WPI;	1992-433371/52.
XX		
XX	Synergistic compsn.	for generating mammalian bone growth -
XX	comprises vitamin-D cpd.	and bone morphogenetic protein
XX		
XX	Disclosure;	Page 31-32; 44pp: English.
XX		
XX	The sequences given	in AAQ32850-56 encode bone morphogenetic proteins
XX	(BMP); BMP's	increase bone growth and when used in conjunction with
XX	vitamin D	the level of new bone growth is greater than when a BMP or
XX	vitamin D	are used alone. The BMP's are administered for systemic
XX	treatment	at a dose range of 1pg to 100 microg. BMP are active as
XX	dimers.	
XX		
XX	Sequence	1751 bp: 394 A: 510 C: 490 G: 357 T: 0 other;
XX		

Query Match 6.2%; Score 22; DB 13; Length 1751;  
Best Local Similarity 100.0%; Pred. No. 0.52;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 ggaggaggggaggggaagga 347  
||| ||| ||| ||| ||| ||| |||  
pb 7 gaagagggagggaggggaagga 28

RESULT 34  
AAV01680  
ID AAV01680 standard: CDNA: 1751 BP.

XX  
AC AAV01680;

XX  
DT 27-MAR-1998 (first entry)

XX DE Bone morphogenetic protein BMP-4 encoding DNA.

xx Bone morphogenetic protein; BMP; growth; vitamin D; fracture;  
kw arthritis; surgical lesion; periodontal disease; osteoporosis;  
kw rickets; ds.

Unidentified.

XX  
PN  
IIS5693615-A.XX  
PD  
02-DEC-1997

XX 05-JUN-1991: 91US-0709621.

XX  
07-SEP-1993. 93US-0117367.

PR 05-JUN-1991; 9TUS=07096ZT.  
PR 27-MAR-1993: 9ZUS-0856110

PR 09-DEC-1992; 9205-0988363.  
PR 13-MAY-1994; 94HS-0243435.  
CC

PR 23 - JAN - 1995; 95US-0377292.

PA (PROC ) PROCTER &amp; GAMBLE CO.

PI Stone RL;

DR WPI; 1998-031788/03.

PT Generation

XX  
XX[illegible]

[illegible]

XX A new method has been developed for generating new bone growth in a  
CC mammal. The method comprises administering a bone morphogenetic protein  
CC in combination with a vitamin D compound, where: (a) the bone  
CC morphogenetic protein is BMP-2 and is administered in an amount of  
CC 500-1000 ng in combination with about 6 ng vitamin D compound; or (b)  
CC the bone morphogenetic protein is BMP-4 and is administered in an  
CC amount of about 62.5 ng in combination with about 6 ng vitamin D  
CC compound. The present sequence encodes BMP-4. The method is used for  
CC treating bone defects or disorders, e.g. fractures, surgical lesions,  
CC periodontal disease, osteoporosis, arthritis and rickets.  
XX  
SQ Sequence 1751 BP; 394 A; 510 C; 490 G; 357 T; 0 other;

```

Query Match      6.2%; Score 22; DB 19; Length 1751;
Best Local Similarity 100.0%; Pred. No. 0.52;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy	326	99aggagggagggaagga	347
Dd	7	99aggagggagggaagga	28

RESULT 35  
AAQ53144  
ID AAQ53144 standard; cDNA; 1788 BP.  
v v

XX  
AC      AAQ5314;  
XX  
DT      06-JUN-1994  
         (first entry)

Sequence encoding osteogenic protein CBMP2B.

osteogenic protein; bone; cartilage; matrix; osteoarthritis;  
 repair; vascularisation; mineralisation; differentiation; ss.  
 xx  
 Homo sapiens.  
 OS

XX	key	Location/Qualifiers
FH	CDS	403..1626
FT		/*tag= a
FT		/product= CBMP2B
FT		

XX  
PN  
XX  
PD  
US5266683-A.  
30-NOV-1993

XX  
XX  
PF 08-APR-1988; 88US-0179406.  
XX

PR 08-APR-1988; 88US-0179406.  
PR 15-AUG-1988; 88US-023630.  
PR 23-FEB-1989; 89US-0315342.  
PR 17-OCT-1989; 89US-0423613.

PR 17-OCT-1989; 89US-0423699.  
PR 22-FEB-1990; 90US-0483913.  
PR 20-AUG-1990; 90US-0569920.

07-SEP-1990; 90US-0575865.  
PR  
18-OCT-1990; 90US-0595543.  
PR  
18-OCT-1990; 90US-0600024.  
PR  
04-DEC-1990; 90US-0621849.  
PR

PR 04-DEC-1990: 90US-0621988.  
PR 22-FEB-1991: 90US-0650162.  
PR 20-DEC-1991: 90US-0810560.  
PR 28-FEB-1992:

K 26 JAN-1992; 92US-0827052.  
 PR 21-FEB-1992; 92US-0841646.  
 XX  
 PI Kuberasampath T, Oppermann H, Ozkavnak E. Pang RHT.

PI  
XX  
Rueger DC;  
WPI: 1993-395405/49.  
DR  
P-PSDB: AAP44749  
DR

XX  
Z. DOL., 2000-2001.

PT	New pure mammalian osteogenic proteins - induce cartilage and
PT	endochondral bone formation when in association with a matrix
XX	
XX	Claim 37; Columns 93-96; 128pp; English.
XX	
CC	The osteogenic protein when in association with a matrix can induce
CC	at the locus of an implant the full development cascade of
CC	endochondral bone formation including vascularisation,
CC	mineralisation and bone marrow differentiation. The osteogenic
CC	protein can also be used to repair both bone and cartilage in the
CC	treatment of osteoarthritis.
XX	
XX	
SQ	Sequence 1788 BP; 418 A; 515 C; 494 G; 361 T; 0 other;

Query Match	6.2%	Score 22;	DB 14;	Length 1788;
Best Local Similarity	100.0%	Pred. No. 0.52;		
Matches 22:	Conservative	0. Mismatches	0. Indels	0. Gaps

[illegible]

RESULT 36  
AAQ72710  
ID AAQ72710 standard. CDNA. 1788 bp

XX  
AC  
AAQ72710;  
XX

DT XX 09-JUN-1995 (first entry)  
DE XX Prepro human CBMP2B cDNA.

Prepro human CBMP2B; cartilage and endochondrial bone formation;  
 KW allograft repair; osteoarthritis; non-union fracture repair;  
 KW osteogenesis; periodontal, dental and craniofacial reconstruction;  
 --

	SS.
KW	
XX	
OS	Homo sapiens.
XX	
XX	
FH	Key
	Location/Qualifiers

```

CDS
403..1629
/*tag= a
XX
US354557-A.

```

XX  
PD 11-OCT-1994.  
XX  
09-SEP-1998.  
XX 0170.105

06-APR-1988;	88US-0179406.
XX	
08-APR-1988;	88US-0179406.
15-AUG-1988;	88US-0232630.

23-FEB-1989; 89US-0315342.  
17-OCT-1989; 89US-0422613.  
17-OCT-1989; 89US-0422699.  
22-FEB-1990; 90US-0493013.

24 JUL 1990; 90US-0463913.  
20-AUG-1990; 90US-0569920.  
07-SEP-1990; 90US-0579865.  
18-OCT-1990; 90US-0599543.

18-OCT-1990; 90US-0600024.  
04-DEC-1990; 90US-0621849.  
04-DEC-1990; 90US-0621888.  
22-FEB-1991; 90US-0650162.

R	20-DEC-1991;	3008 000102.
R	90US-0810560.	
R	28-JAN-1992;	
R	90US-0827052.	
R	21-FEB-1992;	
R	90US-0841646.	

K  
18-DEC-1992; 90US-0993387.  
X  
X  
A  
(STYC ) STRYKER CORP.  
X  
X

I Kuberasampath T, Oppermann H, Ozkaynak E, Pang RHL;  
I Rueger DC;







CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human  
 CC diagnostic coding sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 1832 BP; 447 A; 439 C; 432 G; 484 T; 30 other;

Query Match 6.2%; Score 22; DB 23; Length 1832;  
 Best Local Similarity 100.0%; Pred. No. 0.52;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 328 agagaggaggaggaggaggag 349  
 |||||  
 Db 1321 AGGAGGGAGGGAGGAGGAG 1300

## RESULT 42

AAT78942  
 ID AAT78942 standard; cDNA; 1944 BP.

XX AC AAT78942;

XX DT 22-JAN-1998 (first entry)

XX DE Human bone morphogenic protein (BMP) 2B cDNA.

XX KW bone morphogenic protein; BMP 2B; cartilage; periodontal disease;  
 XX tissue repair; osteoporosis; treatment; oligonucleotide probe; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
 FT 5'UTR 8..402

FT FT /\*tag= a

FT CDS 403..1629

FT FT /\*tag= b

FT mat\_peptide /product= Bone-morphogenic\_protein\_2B

FT FT 1333..1626

FT FT /\*tag= c

FT FT /note= "BMP-2B contains at least this part of the

FT 3'UTR 1630..1938

FT FT /\*tag= d

XX US5631142-A.

XX 20-MAY-1997.

XX 07-SEP-1993; 93US-0118363.

XX 11-JUL-1989; 89US-0378537.

XX 17-DEC-1986; 86US-0943332.

XX 08-MAR-1987; 87US-0028285.

XX 08-APR-1988; 88US-0179100.

XX 18-MAY-1992; 92US-0884353.

XX 07-SEP-1993; 93US-0118363.

XX 01-JUL-1986; 86US-0880776.

XX (GENE ) GENETICS INST INC.

XX Rosen VA, Wang EA, Wozney JM;

XX WPI; 1997-288573/26.

XX P-PSDB; AAN24850.

XX

PT Production of human bone morphogenic protein 2A or 2B in cell  
 PT culture - useful inducing bone or cartilage production, in wound  
 PT healing and tissue repair  
 XX  
 PS Example 5; Fig 3; 22pp; English.  
 XX  
 CC This cDNA sequence encodes the human bone morphogenic protein (BMP) 2B.  
 CC Oligonucleotide probes were synthesised based on a partially sequenced  
 CC BMP-2B protein isolated from ground bovine powder. The probes were used  
 CC to screen a bovine liver DNA library to obtain the BMP-2B encoding DNA  
 CC sequence. The DNA was used to screen a U-2 OS human cell line cDNA  
 CC library to obtain this human BMP-2B cDNA. BMPs can be used to induce  
 CC bone and cartilage formation, and in wound healing and tissue repair.  
 CC They can also be used for treating periodontal disease or osteoporosis.  
 XX  
 SQ Sequence 1944 BP; 488 A; 535 C; 510 G; 411 T; 0 other;

Query Match 6.2%; Score 22; DB 18; Length 1944;  
 Best Local Similarity 100.0%; Pred. No. 0.52;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 ggagggaggaggaggaggaggag 347  
 |||||  
 Db 15 ggagggaggaggaggaggaggag 36

## RESULT 43

AAN80634

ID AAN80634 standard; DNA; 1954 BP.

XX AC AAN80634;

XX DT 08-OCT-1990 (first entry)

XX DE Human Bone Morphogenic Protein-2 class II cDNA from U2OS-3.

XX KW Bone morphogenic protein; hBMP-2 class II; probes; cartilage formation;  
 XX bone formation; osteogenic cpds; prodontal disease; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
 FT CDS 403..1623

FT FT /\*tag= a

FT FT /product=hBMP-2 class II

XX WO8800205-A.

XX 14-JAN-1988.

XX 30-JUN-1987; 87WO-US01537.

XX 26-MAR-1987; 87US-0031346.

XX (GENE-) GENETICS INST INC.

XX Wozney JM, Rosen VA;

XX WPI; 1988-021565/03.

XX P-PSDB; AAP80620.

XX Bone morphogenic proteins - obtd. using recombinant DNA and used  
 XX for inducing cartilage and bone formation.

XX Disclosure; ; 7pp; English.

XX The HindIII-SacI bovine genomic hBMP-2 fragment described in AAN80627  
 XX is subcloned into M13, labelled and used as probe to screen

XX CC polyadenylated RNAs from various cells and tissue sources.

XX CC Sequence analysis of the weakly hybridising recombinants hBMP-2 class II  
 XX (-BMP-4) indicated that they are quite homologous with the sequence

XX given in AAN80622 at the end of their coding regions, but less so in



XX  
SQ Sequence 1954 BP: 492 A: 536

Query Match 6.2%; Score 22; DB 13; Length 1954;  
Best Local Similarity 100.0%; Pred. No. 0.52;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**Oy**            326 ggaggagggagggggaagga 347  
                |  
**Dd**            15 ggaggagggagggggaagga 36

Search completed: September 20, 2002, 06:07:38  
Job time: 10342 sec





```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/377,292
; FILING DATE: 23-JAN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/243,435
; FILING DATE:
; APPLICATION NUMBER: US/08/117,367
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Corstanje, Brahm J.
; REGISTRATION NUMBER: 34,804
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 513-245-2858
; TELEFAX: 513-741-3012
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1751 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-377-292-4

Query Match 6.2%; Score 22; DB 1; Length 1751;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 ggaggaggaggaggaggaggga 347
    |||||
Db 7 GGAGGAGGGAGGGAGGGAAGGA 28

RESULT 3
US-07-841-646-6
; Sequence 6, Application US/07841646
; Patent No. 526683
; GENERAL INFORMATION:
; APPLICANT: OPPERMAN, HERMANN
; APPLICANT: OZKAYNAK, ENGIN
; APPLICANT: KUBERASAMPATH, THANGAVEL
; APPLICANT: RUEGER, DAVID C.
; APPLICANT: PANG, ROY H.L.
; TITLE OF INVENTION: OSTEOGENIC DEVICES
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSER: TESTA, HURWITZ & THIBEAULT
; STREET: 53 STATE STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: U.S.A.
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/841,646
; FILING DATE: 19920221
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 810,560
; FILING DATE: 20-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 827,052
; FILING DATE: 28-JAN-1992

```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 660,162
; FILING DATE: 22-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 621,988
; FILING DATE: 04-DEC-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 621,849
; FILING DATE: 04-DEC-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 616,374
; FILING DATE: 21-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 600,024
; FILING DATE: 18-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 599,543
; FILING DATE: 18-OCT-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 579,865
; FILING DATE: 07-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 569,920
; FILING DATE: 20-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 483,913
; FILING DATE: 22-FEB-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 422,613
; FILING DATE: 17-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 315,342
; FILING DATE: 23-FEB-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 232,630
; FILING DATE: 15-AUG-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 179,460
; FILING DATE: 08-APR-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: PITCHER, EDMUND R.
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: CRP-001CP6
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/248-7000
; TELEFAX: 617/248-7100
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1788 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: HOMO SAPIENS
; TISSUE TYPE: HIPPOCAMPUS
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 403..1626
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: /function= "OSTEOGENIC PROTEIN"
; OTHER INFORMATION: /product= "CBMP2B"
; OTHER INFORMATION: /evidence= EXPERIMENTAL
; OTHER INFORMATION: /note= "CBMP2B (CDNA)"
; US-07-841-646-6

```

```

Query Match 6.2%; Score 22; DB 1; Length 1788;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```



Qy 326 ggaggagggaggggaaaggga 347

Db 15 GGAGGAGGAGGAGGAAGGA 36

## RESULT 7

US-08-449-700-6  
; Sequence 6, Application US/08449700  
; Patent No. 5863758  
; GENERAL INFORMATION:  
; APPLICANT: OPPERMAN, HERMANN  
; APPLICANT: OZKAYNAK, ENGIN  
; APPLICANT: KUBERASAMPATH, THANGAVEL  
; APPLICANT: RUEGER, DAVID C.  
; APPLICANT: PANG, ROY H.L.  
; TITLE OF INVENTION: OSTEOGENIC DEVICES  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT  
; STREET: 53 STATE STREET  
; CITY: BOSTON  
; STATE: MASSACHUSETTS  
; COUNTRY: U.S.A.  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/449,700  
; FILING DATE: 21-FEB-1992  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 810,560  
; FILING DATE: 20-DEC-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 827,052  
; FILING DATE: 28-JAN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 660,162  
; FILING DATE: 22-FEB-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 621,988  
; FILING DATE: 04-DEC-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 621,849  
; FILING DATE: 04-DEC-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 616,374  
; FILING DATE: 21-NOV-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 600,024  
; FILING DATE: 18-OCT-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 599,543  
; FILING DATE: 18-OCT-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 579,865  
; FILING DATE: 07-SEP-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 569,920  
; FILING DATE: 20-AUG-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 483,913  
; FILING DATE: 22-FEB-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 422,613  
; FILING DATE: 17-OCT-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 315,342  
; FILING DATE: 23-FEB-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 232,630  
; FILING DATE: 15-AUG-1988

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 179,460  
; FILING DATE: 08-APR-1988  
; ATTORNEY/AGENT INFORMATION:  
; NAME: PITCHER, EDMUND R.  
; REGISTRATION NUMBER: 27,829  
; REFERENCE/DOCKET NUMBER: CRP-001CP6  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/248-7000  
; TELEFAX: 617/248-7100  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1788 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: HOMO SAPIENS  
; TISSUE TYPE: HIPPOCAMPUS  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 403..1626  
; IDENTIFICATION METHOD: experimental  
; OTHER INFORMATION: /function= "OSTEOGENIC PROTEIN"  
; OTHER INFORMATION: /product= "CBMP2B"  
; OTHER INFORMATION: /evidence= EXPERIMENTAL  
; OTHER INFORMATION: /note= "CBMP2B (CDNA)"  
US-08-449-700-6  
  
Query Match 6.2%; Score 22; DB 2; Length 1788;  
Best Local Similarity 100.0%; Pred. No. 0.053;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 326 ggagggaggaggaggagga 347  
|||||  
Db 15 GGAGGAGGAGGAGGAAGGA 36  
  
RESULT 8  
US-08-449-699A-6  
; Sequence 6, Application US/08449699A  
; Patent No. 5958441  
; GENERAL INFORMATION:  
; APPLICANT: OPPERMAN, HERMANN  
; APPLICANT: OZKAYNAK, ENGIN  
; APPLICANT: KUBERASAMPATH, THANGAVEL  
; APPLICANT: RUEGER, DAVID C.  
; APPLICANT: PANG, ROY H.L.  
; TITLE OF INVENTION: ANTIBODIES TO OSTEOGENIC PROTEINS  
; NUMBER OF SEQUENCES: 33  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT  
; STREET: 125 HIGH STREET  
; CITY: BOSTON  
; STATE: MASSACHUSETTS  
; COUNTRY: U.S.A.  
; ZIP: 02110  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/449,699A  
; FILING DATE: 24-MAY-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/147,023  
; FILING DATE: 01-NOV-1993

ADDRESS: Genetics Institute, Inc.  
 STREET: Legal Affairs - 87 CambridgePark Drive  
 CITY: Cambridge  
 STATE: MA  
 COUNTRY: US  
 ZIP: 02140  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/050,132A  
 FILING DATE:  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Kapinos, Ellen J.  
 REGISTRATION NUMBER: 32,245  
 REFERENCE/DOCKET NUMBER: GI 5186A  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 876-1170  
 TELEFAX: (617) 876-5851  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1954 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA to mRNA  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Homo sapiens  
 CELL TYPE: Osteosarcoma Cell Line  
 CELL LINE: U-20S

Query Match 6.28; Score 22; DB 2; Length 1954;

NAME/KEY:	MRN
LOCATION:	9.





[illegible]

Query Match 6.2%; Score 22; DB 5; Length 1954;  
Best Local Similarity 100.0%; Pred. No. 0.053;

**Qy**            326 ggaggaggagggaagga 347  
                | | | | | | | | | |  
**Db**            15 GGAGCAGGAGGGGAAGCA 36

```

RESULT 19
PCT-US95-07084-3
; Sequence 3, Application PC/TUS9507084
; GENERAL INFORMATION:
; APPLICANT: Rosen, Vicki A.
; APPLICANT: Wozney, John M.
; APPLICANT: Celeste, Anthony J.
; TITLE OF INVENTION: BMP-9 COMPOSITIONS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:

```

```

RESULT 19
PCT-US95-07084-3
; Sequence 3, Application PC/TUS9507084
; GENERAL INFORMATION:
; APPLICANT: Rosen, Vicki A.
; APPLICANT: Wozney, John M.
; APPLICANT: Celeste, Anthony J.
; TITLE OF INVENTION: BMP-9 COMPOSITIONS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
;

```



Db 340 GGAGGGAGGAGGAAGGAAG 320

## RESULT 22

US-08-332-766A-17/C  
; Sequence 17, Application US/08332766A  
; Patent No. 5843647  
; GENERAL INFORMATION:  
; APPLICANT: JEFFREYS, Alec J.  
; APPLICANT: ARMOUR, John  
; TITLE OF INVENTION: SIMPLE TANDEM REPEATS  
; NUMBER OF SEQUENCES: 125  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.  
; STREET: 1100 New York Avenue, N.W.  
; CITY: Washington  
; STATE: D. C.  
; COUNTRY: U.S.A.  
; ZIP: 20005-3918  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/332,766A  
; FILING DATE: 01-NOV-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9326052.9  
; FILING DATE: 21-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BIRD, Donald J.  
; REGISTRATION NUMBER: 25,323  
; REFERENCE/DOCKET NUMBER: 217211/M94/0434/GB  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 861-3000  
; TELEFAX: (202) 822-0944  
; TELEX: 6714627 CUSH  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 591 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; US-08-332-766A-17

Query Match 5.9%; Score 21; DB 2; Length 591;  
Best Local Similarity 100.0%; Pred. No. 0.17;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 329 ggagggagggaggaaggaag 349  
Db 503 GGAGGGAGGAGGAAGGAAG 483

## RESULT 23

US-08-332-766A-42  
; Sequence 42, Application US/08332766A  
; Patent No. 5843647  
; GENERAL INFORMATION:  
; APPLICANT: JEFFREYS, Alec J.  
; APPLICANT: ARMOUR, John  
; TITLE OF INVENTION: SIMPLE TANDEM REPEATS  
; NUMBER OF SEQUENCES: 125  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.  
; STREET: 1100 New York Avenue, N.W.  
; CITY: Washington  
; STATE: D. C.  
; COUNTRY: U.S.A.

; ZIP: 20005-3918  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/332,766A  
; FILING DATE: 01-NOV-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9326052.9  
; FILING DATE: 21-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BIRD, Donald J.  
; REGISTRATION NUMBER: 25,323  
; REFERENCE/DOCKET NUMBER: 217211/M94/0434/GB  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 861-3000  
; TELEFAX: (202) 822-0944  
; TELEX: 6714627 CUSH  
; INFORMATION FOR SEQ ID NO: 42:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 278 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; US-08-332-766A-42

Query Match 5.6%; Score 20; DB 2; Length 278;  
Best Local Similarity 100.0%; Pred. No. 0.52;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 328 agagggagggaggaagga 347  
Db 196 AGGAGGGAGGGAAGGA 215

## RESULT 24

US-09-268-992-7  
; Sequence 7, Application US/09268992  
; Patent No. 6342351  
; GENERAL INFORMATION:  
; APPLICANT: Chen, H.  
; APPLICANT: Freilmer, N.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING  
; TITLE OF INVENTION: AND TREATING CHROMOSOME-18p RELATED DISORDERS  
; FILE REFERENCE: 7853-138  
; CURRENT APPLICATION NUMBER: US/09/268,992  
; CURRENT FILING DATE: 1999-03-16  
; EARLIER APPLICATION NUMBER: 09/236,134  
; EARLIER FILING DATE: 1999-01-22  
; EARLIER APPLICATION NUMBER: 60/106,056  
; EARLIER FILING DATE: 1998-10-28  
; EARLIER APPLICATION NUMBER: 60/088,312  
; EARLIER FILING DATE: 1998-06-05  
; EARLIER APPLICATION NUMBER: 60/078,044  
; EARLIER FILING DATE: 1998-03-16  
; NUMBER OF SEQ ID NOS: 84  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 7  
; LENGTH: 72604  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: all n positions  
; OTHER INFORMATION: n=a, c, g, or t  
US-09-268-992-7

```

RESULT 26
US-08-332-766A-22
; Sequence 22, Application US/08332766A
; Patent No. 5843647
; GENERAL INFORMATION:
; APPLICANT: JEFFREYS, Alec J.
; APPLICANT: ARMOUR, John
; TITLE OF INVENTION: SIMPLE TANDEM
; NUMBER OF SEQUENCES: 125

```

```

27
US-08-713-000-9
Sequence 9, Application US/08713000
Patent No. 5850020
GENERAL INFORMATION:
APPLICANT: Bloksberg, Leonard N.
APPLICANT: Havukkala, Ilkka
APPLICANT: Grierson, Alastair
TITLE OF INVENTION: MATERIALS AND METHODS FOR THE
MODIFICATION OF PLANT LIGNIN CONTENT
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Speckman Picard PLLC
STREET: 2601 Elliott Avenue, Suite 4185
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/713,000
FILING DATE:
CLASSIFICATION: 435

```



; TELEFAX: 206-269-0563  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 45:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 684 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-975-316-45

Query Match 5.3%; Score 19; DB 2; Length 684;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 gctggcgctgctgctgag 187  
|||||  
Db 261 GCTGGCGTCTGCTGAG 279

RESULT 31  
US-08-705-771-8  
; Sequence 8, Application US/08705771  
; Patent No. 6054289  
; GENERAL INFORMATION:  
; APPLICANT: Paul Moore, Reiner Gentz, Hongjin Ji,  
; APPLICANT: Jian Ni and Jing-Shan Hu  
; TITLE OF INVENTION: Human Genes, Sequences and  
; TITLE OF INVENTION: Expression Products  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
; ADDRESSEE: CECCHI, STEWART & OLSTEIN  
; STREET: 6 BECKER FARM ROAD  
; CITY: ROSELAND  
; STATE: NEW JERSEY  
; COUNTRY: USA  
; ZIP: 07068  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 INCH DISKETTE  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WORD PERFECT 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/705,771  
; FILING DATE: August 30, 1996  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MULLINS, J.G.  
; REGISTRATION NUMBER: 33,073  
; REFERENCE/DOCKET NUMBER: 325800-346 (PF196)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 973-994-1700  
; TELEFAX: 973-994-1744  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1344 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA  
US-08-705-771-8

Query Match 5.3%; Score 19; DB 3; Length 1344;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 ggaaggaggaggaggaa 344  
|||||  
Db 959 GGAGGAGGAGGAGGGAA 977

RESULT 32  
US-09-370-807-7/c  
; Sequence 7, Application US/09370807  
; Patent No. 6297034  
; GENERAL INFORMATION:  
; APPLICANT: Cahoon, Rebecca E.  
; APPLICANT: Falco, S. Carl  
; APPLICANT: Rafalski, J. Antoni  
; APPLICANT: Sakai, Hajime  
; TITLE OF INVENTION: N-End Rule Pathway Enzymes  
; FILE REFERENCE: BB-1199  
; CURRENT APPLICATION NUMBER: US/09/370,807  
; CURRENT FILING DATE: 1999-08-09  
; EARLIER APPLICATION NUMBER: 60/096,225  
; EARLIER FILING DATE: August 12, 1998  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 7  
; LENGTH: 2407  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
US-09-370-807-7

Query Match 5.0%; Score 18; DB 4; Length 2407;  
Best Local Similarity 100.0%; Pred. No. 5.5;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 gcttgcagcaataactg 67  
|||||  
Db 513 GCTTGCAGCAATAACTG 496

RESULT 33  
US-08-765-662-13  
; Sequence 13, Application US/08765662  
; Patent No. 5929213  
; GENERAL INFORMATION:  
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE  
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-12  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/765,662  
; FILING DATE: 28-APR-1997  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/08745  
; FILING DATE: 12-JUL-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Haile, Ph.D., Lisa A  
; REGISTRATION NUMBER: 38,347  
; REFERENCE/DOCKET NUMBER: 07265/042W01 (FD-3830)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-678-5070  
; TELEFAX: 619-678-5099  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2419 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single

; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; FEATURE:  
; NAME/KEY: Coding Sequence  
; LOCATION: 218...1267  
; OTHER INFORMATION:  
US-08-765-662-13

Query Match 5.0%; Score 18; DB 2; Length 2419;  
Best Local Similarity 100.0%; Pred. No. 5.5;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 ggaggaggaggaggagg 346  
|||||  
Db 1563 GGAGGAGGAGGAGGAGG 1580

## RESULT 34

PCT-US95-08745-13  
; Sequence 13, Application PC/TUS9508745  
; GENERAL INFORMATION:  
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE  
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-12  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/08745  
; FILING DATE: 12-JUL-1995  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Haile, Ph.D., Lisa A  
; REGISTRATION NUMBER: 38,347  
; REFERENCE/DOCKET NUMBER: 07265/042W01 (FD-3830)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-678-5070  
; TELEFAX: 619-678-5099  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2419 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE:  
; ORIGINAL SOURCE:  
; FEATURE:  
; NAME/KEY: Coding Sequence  
; LOCATION: 218...1267  
; OTHER INFORMATION:  
PCT-US95-08745-13

Query Match 5.0%; Score 18; DB 5; Length 2419;  
Best Local Similarity 100.0%; Pred. No. 5.5;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 329 ggaggaggaggaggagg 346  
|||||  
Db 1563 GGAGGAGGAGGAGGAGG 1580

## RESULT 35

US-09-173-914-1  
; Sequence 1, Application US/09173914  
; Patent No. 6171857  
; GENERAL INFORMATION:  
; APPLICANT: Hendrickson, Eric  
; TITLE OF INVENTION: A No. 6171857el Leucine zipper, KARP-1 and  
; FILE OF INVENTION: Methods of Regulating DNA Dependent Protein Kinase Activity  
; FILE REFERENCE: 80877/7017/HK  
; CURRENT APPLICATION NUMBER: US/09/173,914  
; CURRENT FILING DATE: 1998-10-16  
; EARLIER APPLICATION NUMBER: 60/064,557  
; EARLIER FILING DATE: 1997-10-17  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 6078  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (178)...(178)  
; NAME/KEY: unsure  
; LOCATION: (230)...(230)  
; NAME/KEY: unsure  
; LOCATION: (232)...(232)  
; NAME/KEY: unsure  
; LOCATION: (234)...(234)  
; NAME/KEY: unsure  
; LOCATION: (453)...(453)  
; NAME/KEY: unsure  
; LOCATION: (473)...(473)  
; NAME/KEY: unsure  
; LOCATION: (610)...(610)  
; NAME/KEY: unsure  
; LOCATION: (612)...(612)  
; NAME/KEY: unsure  
; LOCATION: (2175)...(2175)  
; NAME/KEY: unsure  
; LOCATION: (1014)...(1014)  
US-09-173-914-1

Query Match 5.0%; Score 18; DB 4; Length 6078;  
Best Local Similarity 100.0%; Pred. No. 5.6;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 330 ggaggaggaggaggagg 347  
|||||  
Db 204 ggaggaggaggaggagg 221

## RESULT 36

US-08-483-376-1/c  
; Sequence 1, Application US/08483376  
; Patent No. 5955330  
; GENERAL INFORMATION:  
; APPLICANT: Vasil, Vimla  
; APPLICANT: Clancy, Maureen A.  
; APPLICANT: Ferl, Robert J.  
; APPLICANT: Vasil, Indra K.  
; APPLICANT: Hannah, L. C.  
; TITLE OF INVENTION: No. 5955330el Means for Enhancing Gene  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:



ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
STREET: 5370 Manhattan Circle, Suite 201  
CITY: Boulder  
STATE: Colorado  
COUNTRY: US  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US 08/483,376  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/418,540  
FILING DATE: 07-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/102,115  
FILING DATE: 04-AUG-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/830,956  
FILING DATE: 05-FEB-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/353,854  
FILING DATE: 18-MAY-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Ferber, Donna M.  
REGISTRATION NUMBER: 33,878  
REFERENCE/DOCKET NUMBER: 10-94B  
TELEPHONE: (303) 499-8080  
TELEFAX: (303) 499-8089  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6386 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Zea mays  
STRAIN: Black Sweet  
FEATURE:  
NAME/KEY: exon  
LOCATION: 131..182  
FEATURE:  
NAME/KEY: exon  
LOCATION: 1211..1324  
FEATURE:  
NAME/KEY: exon  
LOCATION: 1828..1948  
FEATURE:  
NAME/KEY: exon  
LOCATION: 2041..2187  
FEATURE:  
NAME/KEY: exon  
LOCATION: 2269..2460  
FEATURE:  
NAME/KEY: exon  
LOCATION: 2605..2728  
FEATURE:  
NAME/KEY: exon  
LOCATION: 2822..3038  
FEATURE:  
NAME/KEY: exon  
LOCATION: 3256..3351  
FEATURE:  
NAME/KEY: exon  
LOCATION: 3447..3620

FEATURE:  
NAME/KEY: exon  
LOCATION: 3702..3818  
FEATURE:  
NAME/KEY: exon  
LOCATION: 3912..4078  
FEATURE:  
NAME/KEY: exon  
LOCATION: 4158..4381  
FEATURE:  
NAME/KEY: exon  
LOCATION: 4517..4835  
FEATURE:  
NAME/KEY: exon  
LOCATION: 4768..5212  
FEATURE:  
NAME/KEY: exon  
LOCATION: 5372..5510  
FEATURE:  
NAME/KEY: exon  
LOCATION: 5636..5917  
US-08-483-376-1  
Query Match 5.0%; Score 18; DB 2: Length 6386;  
Best Local Similarity 100.0%; Pred. No. 5.6;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 325 tggaggaggaggaggagg 342  
|||||  
Db 151 TGGAGGAGGGAGGGAGG 134  
RESULT 37  
US-08-814-095-7/c  
Sequence 7, Application US/08814095  
Patent No. 6025183  
GENERAL INFORMATION:  
APPLICANT: Soreq, Hermona  
APPLICANT: Zakut, Haim  
APPLICANT: Shani, Moshe  
TITLE OF INVENTION: TRANSGENIC ANIMAL ASSAY SYSTEM FOR  
TITLE OF INVENTION: ANTI-CHOLINESTERASE SUBSTANCES  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: KOHN & ASSOCIATES  
STREET: 30500 No. 6025183thwestern Highway, Suite 410  
CITY: Farmington Hills  
STATE: Michigan  
COUNTRY: U.S.  
ZIP: 48334  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/814,095  
FILING DATE:  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Montgomery, Ilene N.  
REGISTRATION NUMBER: 38,972  
REFERENCE/DOCKET NUMBER: 2391.00066  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (248) 539-5050  
TELEFAX: (248) 539-5055  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35060 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear

```
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Cosmid including ACHE
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
POSITION IN GENOME:
CHROMOSOME/SEGMENT: 7q22
FEATURE:
NAME/KEY: promoter
LOCATION: 4089..22464
OTHER INFORMATION: /function= "ACHE Promotor"
OTHER INFORMATION: /standard_name= "ACHE Promotor"
FEATURE:
NAME/KEY: exon
LOCATION: 22465..22537
OTHER INFORMATION: /function= "non-translated"
OTHER INFORMATION: /gene= "ACHE"
OTHER INFORMATION: /number= 1
FEATURE:
NAME/KEY: exon
LOCATION: 24090..25177
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /function= "(translation start:
OTHER INFORMATION: 24110)"
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "ACHE"
OTHER INFORMATION: /number= 2
FEATURE:
NAME/KEY: exon
LOCATION: 25524..26009
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "ACHE"
OTHER INFORMATION: /number= 3
FEATURE:
NAME/KEY: exon
LOCATION: 27005..27274
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "ACHE"
OTHER INFORMATION: /number= 4
FEATURE:
NAME/KEY: exon
LOCATION: 27255..28007
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "ACHE"
OTHER INFORMATION: /number= 5
FEATURE:
NAME/KEY: terminator
LOCATION: 27385..27387
FEATURE:
NAME/KEY: exon
LOCATION: 28008..28129
IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /evidence= EXPERIMENTAL
OTHER INFORMATION: /gene= "ACHE"
OTHER INFORMATION: /number= 6
FEATURE:
NAME/KEY: terminator
LOCATION: 28129..28131
FEATURE:
NAME/KEY: exon
LOCATION: complement (34528..34895)
OTHER INFORMATION: /function= "arsenite resistance
OTHER INFORMATION: gene"
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 1
FEATURE:
NAME/KEY: exon
LOCATION: complement (34092..34358)
```

```
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 2
FEATURE:
NAME/KEY: exon
LOCATION: complement (33779..33963)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 3
FEATURE:
NAME/KEY: exon
LOCATION: complement (33493..33591)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 4
FEATURE:
NAME/KEY: exon
LOCATION: complement (33297..33408)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 5
FEATURE:
NAME/KEY: exon
LOCATION: complement (32959..33094)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 6
FEATURE:
NAME/KEY: exon
LOCATION: complement (32569..32628)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 7
FEATURE:
NAME/KEY: exon
LOCATION: complement (32386..32468)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 8
FEATURE:
NAME/KEY: exon
LOCATION: complement (31894..32080)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 9
FEATURE:
NAME/KEY: exon
LOCATION: complement (31363..31534)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 10
FEATURE:
NAME/KEY: exon
LOCATION: complement (31131..31284)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 11
FEATURE:
NAME/KEY: exon
LOCATION: complement (30816..31011)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 12
FEATURE:
NAME/KEY: exon
LOCATION: complement (30470..30626)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 13
FEATURE:
NAME/KEY: exon
LOCATION: complement (30187..30274)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 14
FEATURE:
NAME/KEY: exon
LOCATION: complement (29945..30073)
OTHER INFORMATION: /gene= "AR"
OTHER INFORMATION: /number= 15
FEATURE:
NAME/KEY: exon
LOCATION: complement (29664..29856)
OTHER INFORMATION: /gene= "ARS"
OTHER INFORMATION: /number= 16
US-08-814-095-7
```

APPLICANT: CHENIVESSE, Xavier  
APPLICANT: HENGLEIN, Berthold  
APPLICANT: ZINDY, Fr d rique  
TITLE OF INVENTION: NEW HUMAN CYCLIN A COMPOSITIONS AND A PROCESS FOR THEIR PRO  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Larson and Taylor  
STREET: 727 Twenty-Third Street, South  
CITY: Arlington  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette-5.25 inch, 500 Kb storage  
COMPUTER: IBM PC/XT/AT or compatibles  
OPERATING SYSTEM: MS-DOS version 3.0 or above  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,895  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/650, 805  
FILING DATE: 06-FEB-1991  
APPLICATION NUMBER: FR9001596  
FILING DATE: 12-FEB-1990  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1634 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cdna  
ORGANISM: HOMO SAPIENS  
FEATURE:  
NAME/KEY: Human cyclin A  
LOCATION: coding sequence from base 97 to base 1392,  
LOCATION: coding for a protein of 432 amino acids.  
US-08-460-895-1  
Query Match 4.8%; Score 17; DB 2; Length 1634;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 56 cagcaataactgaatgc 72  
Db 1165 CAGCAATAACTGATGCG 1149  
RESULT 40  
US-09-210-889-1/c  
Sequence 1, Application US/09210889  
Patent No. 6103887  
GENERAL INFORMATION:  
APPLICANT: BRECHOT, Christian  
APPLICANT: WANG, Jian  
APPLICANT: CHENIVESSE, Xavier  
APPLICANT: HENGLEIN, Berthold  
APPLICANT: ZINDY, Fr d rique  
TITLE OF INVENTION: NEW HUMAN CYCLIN A COMPOSITIONS AND A  
TITLE OF INVENTION: PROCESS FOR THEIR PRODUCTION, THE CORRESPONDING NUCLEOTIDE  
TITLE OF INVENTION: A PROCESS AND AGENTS FOR CELL PROLIFERATION DETECTION OR DI  
TITLE OF INVENTION: AND A PROCESS AND AGENTS FOR INHIBITING CELL PROLIFERATION.  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Larson and Taylor  
STREET: 727 Twenty-Third Street, South  
CITY: Arlington  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22202

Query Match 5.0%; Score 18; DB 3; Length 35060;  
Best Local Similarity 100.0%; Pred. No. 5.7;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 330 gagggaggagggaagga 347  
Db 23954 GAGGGAGGAGGGAAGGA 23937  
RESULT 38  
US-08-332-766A-4/c  
Sequence 4, Application US/08332766A  
Patent No. 5843647  
GENERAL INFORMATION:  
APPLICANT: JEFFREYS, Alec J.  
APPLICANT: ARMOUR, John  
TITLE OF INVENTION: SIMPLE TANDEM REPEATS  
NUMBER OF SEQUENCES: 125  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.  
STREET: 1100 New York Avenue, N.W.  
CITY: Washington  
STATE: D. C.  
COUNTRY: U.S.A.  
ZIP: 20005-3918  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/332,766A  
FILING DATE: 01-NOV-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9326052.9  
FILING DATE: 21-DEC-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: BIRD, Donald J.  
REGISTRATION NUMBER: 25,323  
REFERENCE/DOCKET NUMBER: 217211/M94/0434/GB  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 861-3000  
TELEFAX: (202) 822-0944  
TELEX: 6714627 CUSH  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 217 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-332-766A-4  
Query Match 4.8%; Score 17; DB 2; Length 217;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 331 agggaggagggaagga 347  
Db 102 AGGGAGGAGGGAAGGA 86  
RESULT 39  
US-08-460-895-1/c  
Sequence 1, Application US/08460895  
Patent No. 5849508  
GENERAL INFORMATION:  
APPLICANT: BRECHOT, Christian  
APPLICANT: WANG, Jian

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette-5.25 inch, 500 Kb storage  
COMPUTER: IBM PC/XT/AT or compatibles  
OPERATING SYSTEM: MS-DOS version 3.0 or above  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/210,889  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/650, 805  
FILING DATE: 06-FEB-1991  
APPLICATION NUMBER: FR9001596  
FILING DATE: 12-FEB-1990  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1634 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: HOMO SAPIENS  
FEATURE:  
NAME/KEY: Human cyclin A  
LOCATION: coding sequence from base 97 to base 1392,  
LOCATION: coding for a protein of 432 amino acids.  
US-09-210-889-1

Query Match 4.8%; Score 17; DB 3; Length 1634;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 cagcaataactgatgccc 72  
|||||

Db 1165 CAGCAATAACTGATGCG 1149

RESULT 41  
US-08-692-787-8/c  
Sequence 8, Application US/08692787  
Patent No. 5882864  
GENERAL INFORMATION:  
APPLICANT: An, Gang  
APPLICANT: O'Hara, S. Mark  
APPLICANT: Ralph, David  
APPLICANT: Veltri, Robert  
TITLE OF INVENTION: BIOMARKERS AND TARGETS FOR DIAGNOSIS,  
TITLE OF INVENTION: PROGNOSIS AND MANAGEMENT OF PROSTATE  
DISEASE  
NUMBER OF SEQUENCES: 82  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: United States of America  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/692,787  
FILING DATE: Concurrently Herewith  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Corder, Timothy S.  
REGISTRATION NUMBER: 38,414  
REFERENCE/DOCKET NUMBER: UROC-012  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (512) 418-3000  
TELEFAX: (512) 474-7577  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1649 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-692-787-8

Query Match 4.8%; Score 17; DB 2; Length 1649;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 cagcaataactgatgccc 72  
|||||

Db 1176 CAGCAATAACTGATGCG 1160

RESULT 42  
US-09-097-199-8/c  
Sequence 8, Application US/09097199  
Patent No. 6218529  
GENERAL INFORMATION:  
APPLICANT: An, Gang  
APPLICANT: O'Hara, S. Mark  
APPLICANT: Ralph, David  
APPLICANT: Veltri, Robert  
TITLE OF INVENTION: BIOMARKERS AND TARGETS FOR DIAGNOSIS,  
TITLE OF INVENTION: PROGNOSIS AND MANAGEMENT OF PROSTATE DISEASE  
NUMBER OF SEQUENCES: 87  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/097,199  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/692,787  
FILING DATE: 31-JUL-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Nakashima, Richard A.  
REGISTRATION NUMBER: P-42,023  
REFERENCE/DOCKET NUMBER: UROC:018  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (512) 474-7577  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1649 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-097-199-8

Query Match 4.8%; Score 17; DB 4; Length 1649;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 cagcaataactgatgccc 72  
|||||

Db 1176 CAGCAATACTGATGGC 1160

## RESULT 43

US-09-078-294-4/c  
; Sequence 4, Application US/09078294  
; Patent No. 6265211  
; GENERAL INFORMATION:  
; APPLICANT: Choo, Kong-Hong Andy  
; APPLICANT: Du Sart, Desiree  
; APPLICANT: Cancilla, Michael R.  
; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE  
; FILE REFERENCE: Davies Col  
; CURRENT APPLICATION NUMBER: US/09/078,294  
; CURRENT FILING DATE: 1998-05-13  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 80246  
; TYPE: DNA  
; ORGANISM: Nucleotide sequence of NC-contig  
US-09-078-294-4

Query Match 4.8%; Score 17; DB 4; Length 80246;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 ggaggaggaggagggaag 345  
|||||  
Db 21481 GGAGGGAGGGAGGAAG 21465

## RESULT 44

US-09-078-294-3/c  
; Sequence 3, Application US/09078294  
; Patent No. 6265211  
; GENERAL INFORMATION:  
; APPLICANT: Choo, Kong-Hong Andy  
; APPLICANT: Du Sart, Desiree  
; APPLICANT: Cancilla, Michael R.  
; TITLE OF INVENTION: A NOVEL NUCLEIC ACID MOLECULE  
; FILE REFERENCE: Davies Col  
; CURRENT APPLICATION NUMBER: US/09/078,294  
; CURRENT FILING DATE: 1998-05-13  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 80595  
; TYPE: DNA  
; ORGANISM: Nucleotide sequence of HC-contig  
US-09-078-294-3

Query Match 4.8%; Score 17; DB 4; Length 80595;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 ggaggaggaggagggaag 345  
|||||  
Db 21743 GGAGGGAGGGAGGAAG 21727

## RESULT 45

US-09-109-663-14  
; Sequence 14, Application US/09109663  
; Patent No. 6277981  
; GENERAL INFORMATION:  
; APPLICANT: Tu, Guang-Chou  
; APPLICANT: Israel, Yedy  
; TITLE OF INVENTION: AN IMPROVED METHOD FOR DESIGN AND SELECTION OF  
; TITLE OF INVENTION: EFFICACIOUS ANTISENSE OLIGONUCLEOTIDES  
; FILE REFERENCE: 9855-301

; CURRENT APPLICATION NUMBER: US/09/109,663  
; CURRENT FILING DATE: 1998-07-03  
; EARLIER APPLICATION NUMBER: 60/051,705  
; EARLIER FILING DATE: 1997-07-03  
; NUMBER OF SEQ ID NOS: 81  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 14  
; LENGTH: 21  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Candidate  
; OTHER INFORMATION: TNF(alpha) ASO  
US-09-109-663-14

Query Match 4.5%; Score 16; DB 4; Length 21;  
Best Local Similarity 100.0%; Pred. No. 51;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 334 gagggagggaaggaag 349  
|||||  
Db 3 gagggagggaaggaag 18

Search completed: September 20, 2002, 06:15:23  
Job time: 11142 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 04:07:18 ; Search time 3900.56 Seconds  
(without alignments)  
1235.313 Million cell updates/sec

Title: US-09-846-456-2  
Perfect score: 357  
Sequence: 1 ttgaggtctcactgagagg.....gagggaaggaactgtgttg 357

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0  
Searched: 13736207 seqs, 6748477542 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_estl:\*  
10: gb\_estd:\*  
11: gb\_hic:\*  
12: gb\_gss:\*  
13: em\_gss\_hum:\*  
14: em\_gss\_inv:\*  
15: em\_gss\_pln:\*  
16: em\_gss\_vrt:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	14.3	736	9 AU135588	AU135588 AU135588
2	39	10.9	535	10 BG384217	BG384217 303216 MA
3	24	6.7	158	12 AQ508395	AQ508395 RPCI-11-2
4	24	6.7	292	10 Z44377	Z44377 HSC12B081 n
5	24	6.7	709	12 AQ528018	AQ528018 RPCI-11-3
6	24	6.7	716	12 AG163359	AG163359 Pan trogl
7	23	6.4	401	9 AT394181	AT394181 tg67a04.x
8	23	6.4	827	12 B20638	B20638 F15016-r7 I
9	22	6.2	289	9 BE050974	BE050974 za71c12.g
10	22	6.2	303	12 BH293527	BH293527 CH230-30L
11	22	6.2	371	10 BF592428	BF592428 7156a03.x
12	22	6.2	406	10 R11136	R11136 yf39b09.r1
13	22	6.2	465	9 A1481405	A1481405 vg17d11.x
14	22	6.2	478	9 BE050973	BE050973 za71c12.b
15	22	6.2	483	12 AZ654841	AZ654841 LM0529P18
16	22	6.2	515	12 BH337839	BH337839 CH230-194
17	22	6.2	523	10 BM310687	BM310687 ig47b12.y

c 18	6.2	22	525	12	AZ647100
c 19	6.2	22	544	12	AQ415718
c 20	6.2	22	552	9	AV938157
c 21	6.2	22	569	12	AQ078910
c 22	6.2	22	572	12	AZ378099
c 23	6.2	22	591	12	AZ828464
c 24	6.2	22	615	12	AQ924975
c 25	6.2	22	623	12	B70930
c 26	6.2	22	673	10	BF984535
c 27	6.2	22	713	12	AZ742461
c 28	6.2	22	754	12	AQ983682
c 29	6.2	22	760	10	BE877424
c 30	6.2	22	907	10	BE617281
c 31	6.2	22	926	10	BG823588
c 32	6.2	22	1068	10	BG328313
c 33	5.9	21	43	12	AZ995124
c 34	5.9	21	51	12	AZ351915
c 35	5.9	21	62	12	AZ500262
c 36	5.9	21	66	12	AZ411857
c 37	5.9	21	90	12	AZ838528
c 38	5.9	21	120	12	AZ743191
c 39	5.9	21	145	9	AI053852
c 40	5.9	21	161	12	AZ091117
c 41	5.9	21	191	12	AZ457379
c 42	5.9	21	212	12	AZ656509
c 43	5.9	21	269	10	F11105
c 44	5.9	21	294	12	AZ513751
c 45	5.9	21	319	12	B50482

## ALIGNMENTS

RESULT 1  
AU135588  
LOCUS AU135588 PLACEL Homo sapiens cDNA clone PLACE1002437 5', mRNA  
DEFINITION AU135588 736 bp  
ACCESSION AU135588.1 GI:10996127  
VERSION AU135588.1  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 736)  
AUTHORS Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and Isogai, T.  
TITLE HRI human cDNA project  
JOURNAL Unpublished (2000)  
COMMENT Contact: Takao Isogai  
Genomics Laboratory  
Helix Research Institute  
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
Tel: 81-438-52-3951  
Fax: 81-438-52-3952  
Email: genomics@hri.co.jp  
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.  
FEATURES  
Location/Qualifiers  
1..736  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="PLACE1002437"  
/clone\_lib="PLACE1"  
/tissue\_type="placenta"  
/note="Vector: pME18SFL3"

BASE COUNT 163 a 199 c 199 g 170 t 5 others  
ORIGIN

Query Match 14.3%; Score 51; DB 9; Length 736;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-14;  
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 208 cagctgaggttgcctgtctggaagaacctcttccagagaagacaaca 258  
 |||||||  
 Db 329 CAGCTGAGGTTGCTGCTGTGGAAGAACCTCACTTTTCAGAAGACAACA 379  
 |||||||

RESULT 2  
 BG384217  
 LOCUS 303216 MARC LP1G Sus scrofa cDNA 5', mRNA linear EST 12-MAR-2001  
 ACCESSION BG384217  
 VERSION BG384217.1 GI:13308689  
 KEYWORDS EST.  
 SOURCE pig.  
 ORGANISM Sus scrofa  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 REFERENCE 1 (bases 1 to 535)  
 AUTHORS Fahrenkrug, S.C., Preking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E.,  
 Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Laegreid, W.W.  
 and Keefe, J.W.  
 TITLE Design and use of two pooled tissue normalized cDNA libraries for  
 EST discovery in swine  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Smith TPL  
 USDA, ARS, US Meat Animal Research Center  
 PO Box 166, Clay Center, NE 68933-0166, USA  
 Tel: 402 762 4366  
 Fax: 402 762 4390  
 Email: smith@email.marc.usda.gov  
 Single pass sequencing. Bases called and alt\_trimmed with phred  
 v0.980904.e. Vector identified by cross\_match with the -minscore 18  
 and -minmatch 12 options.  
 PCR Primers  
 FORWARD: AGGAACAGCATATGACCAT  
 BACKWARD: GTTTCCTCCAGTCACGACG  
 Plate: 90 row: G column: 13  
 Seq primer: ATTTAGTGACACTATAG.  
 FEATURES source  
 Location/Qualifiers  
 1..535  
 /organism="Sus scrofa"  
 /db\_xref="taxon:9823"  
 /clone\_lib="MARC LP1G"  
 /tissue\_type="pooled"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV SPORT6; Site\_1: XbaI; Site\_2: XhoI;  
 Library made from pooled tissue from day 11, 13, 15, 20,  
 and 30 embryos."  
 BASE COUNT 121 a 159 c 136 g 119 t

Query Match 10.9%; Score 39; DB 10; Length 535;  
 Best Local Similarity 100.0%; Pred. No. 1.6e-08;  
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 ctgctggaagaacctcttccagagaagacaaca 258  
 |||||||  
 Db 311 CTGCTGGAAGAACCTCACTTTTCAGAAGACAACA 349  
 |||||||

RESULT 3  
 AQ508395  
 LOCUS 158 bp DNA linear GSS 29-APR-1999  
 DEFINITION RPCI-11-274E11.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-274E11  
 , DNA sequence.  
 ACCESSION AQ508395  
 VERSION AQ508395.1 GI:4713142  
 KEYWORDS GSS.

SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 158)  
 AUTHORS Zhao, S., Adams, M.D., Niernan, W., Malek, J., de Jong, P. and Venter  
 , J.C.  
 TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready  
 Map Building  
 JOURNAL Unpublished (1997)  
 COMMENT Other\_GSSs: RPCI-11-274E11.TV  
 Contact: Shaying Zhao, William Niernan, Mark Adams  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: hbe@tigr.org  
 Clones are derived from the human BAC library RPCI-11. For BAC  
 library availability, please contact Pieter de Jong  
 (pieter@dejong.med.buffalo.edu). Clones may be purchased from  
 BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from  
 Research Genet cs (info@resgen.com). BAC end search page:  
 http://www.tigr.org/tldb/humgen/bac\_end\_search/bac\_end\_search.html.  
 Seq primer: SP6  
 Class: BAC ends.  
 FEATURES Location/Qualifiers  
 1..158  
 /organism="Homo sapiens"  
 /db\_xref="GDB:7604938"  
 /db\_xref="taxon:9606"  
 /clone="RPCI-11-274E11"  
 /clone\_lib="RPCI-11"  
 /sex="Male"  
 /cell\_type="Lymphocytes"  
 /vector="Vector: pBAC3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
 RPCI11 Human Male BAC Library"  
 BASE COUNT 58 a 12 c 69 g 19 t

Query Match 6.7%; Score 24; DB 12; Length 158;  
 Best Local Similarity 100.0%; Pred. No. 0.37;  
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 ggaggaggaggaggaggaggaggagg 349  
 |||||||  
 Db 120 GGAGGAGGAGGAGGAGGAGGAGGAG 143  
 |||||||

RESULT 4  
 Z44377  
 LOCUS 292 bp mRNA linear EST 14-NOV-1994  
 DEFINITION HSC1ZB081 normalized infant brain cDNA Homo sapiens cDNA clone  
 c-1zb08, mRNA sequence.  
 ACCESSION Z44377  
 VERSION Z44377.1 GI:573506  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 292)  
 AUTHORS Auffray, C., Behar, G., Bois, F., Bouchier, C., da Silva, C., Devignes  
 , M.D., Duprat, S., Houlgatte, R., Jumeau, M.N., Lamy, B., Lorenzo, F.,  
 Mitchell, H., Mariage-Samson, R., Pietu, G., Pouliot, Y.,  
 Sebastiani-Kabaktchis, C. and Tessier, A.  
 TITLE IMAGE: molecular integration of the analysis of the human genome  
 and its expression  
 JOURNAL C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)  
 MEDLINE 95277534  
 COMMENT Contact: Genethon  
 Genexpress-Genethon



Genethon Centre de recherche sur le Genome Humain  
1,rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE  
Tel: 33169472800  
Fax: 33160778698  
Email: genexpres@genethon.fr  
Single read.  
Genexpres library idt: C; Genexpres\_sequence\_idt: ylc-1zb08  
Seq primer: (-21)M13 universal.  
Location/Qualifiers

## FEATURES

source

1..292  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="C-lzb08"  
/clone\_lib="normalized infant brain cDNA"  
/sex="female"  
/tissue\_type="total brain"  
/dev\_stages="3 months old"  
/note="Organ: brain; Vector: lafmid BA; Site\_1: HindIII;  
Site\_2: NotI; sex=female; dev\_stage=3 months old;  
isolate=muscular atrophy patient; tissue\_type=total brain  
; total mRNA was oligo-(dT) primed and directionally  
cloned 5' -> 3' into the HindIII -> NotI sites of the  
lafmid BA vector. Clone library from B.Souares, Psychiatry  
Dept. Columbia University, USA. Normalization\_method:  
Bento Soares, P.N.A.S in press"

BASE COUNT 50 a 87 c 56 t 3 others

ORIGIN

Query Match 6.7%; Score 24; DB 10; Length 292;  
Best Local Similarity 100.0%; Pred. No. 0.4;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 agctctggcgtgcttcacagg 154

|||||

DB 236 AGCTCTGGCGCTGCTTCAGGG 259

RESULT 5  
AQ528018/c

LOCUS 709 bp DNA linear GSS 18-MAY-1999  
DEFINITION RPCI-11-313F19.TV RPCI-11 Homo sapiens genomic clone RPCI-11-313F19  
DNA sequence.

ACCESSION AQ528018.1 GI:4840172

VERSION AQ528018.1

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 709)  
Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter  
J.C.

TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready  
Map Building

JOURNAL Map Building

COMMENT Unpublished (1997)

Other GSSs: RPCI-11-313F19.TJ

Contact: Shaying Zhao, William Nierman, Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850

Tel: 301 838 0200

Fax: 301 838 0208

Email: hbeetigr.org

Clones are derived from the human BAC library RPCI-11. For BAC  
library availability, please contact Pieter de Jong  
(pieter@dejong.med.buffalo.edu). Clones may be purchased from  
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from  
Research Genet cs (info@resgen.com). BAC end search page:

http://www.tigr.org/tcdb/hungen/bac\_end\_search/bac\_end\_search.html.

Seq primer: T7

Class: BAC ends.

Location/Qualifiers

FEATURES

source

1..709  
/organism="Homo sapiens"  
/db\_xref="GDB:7619946"  
/db\_xref="taxon:9606"  
/clone="RPCI-11-313F19"  
/clone\_lib="RPCI-11"  
/sex="Male"  
/cell\_type="Lymphocytes"  
/note="Vector: pBACE3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
RPCI11 Human Male BAC Library"

BASE COUNT 159 a 157 c 100 g 293 t

ORIGIN

Query Match 6.7%; Score 24; DB 12; Length 709;

Best Local Similarity 100.0%; Pred. No. 0.44; 0; Indels 0; Gaps 0;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 ggaggaggaggaggaggaggagg 349

|||||

DB 84 GGAGGAGGAGGAGGAGGAGGAG 61

RESULT 6

AG163359/c

LOCUS 716 bp DNA linear GSS 09-JAN-2002

DEFINITION Pan troglodytes DNA, clone: RP43-030B04.T7, genomic survey

sequence.

ACCESSION AG163359

VERSION AG163359.1 GI:16693037

KEYWORDS GSS (genome survey sequence).

SOURCE Pan troglodytes male lymphocytes DNA, clone\_lib:RPCI-43 Chimpanzee

Male BAC Library clone:RP43-030B04.T7.

ORGANISM Pan troglodytes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

REFERENCE 1 (sites)

AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,

Totoki,Y., Watanabe,H. and Sakaki,Y.

BAC end sequences of Library RPCI-43

Unpublished

REFERENCE 2 (bases 1 to 716)

AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,

Totoki,Y., Watanabe,H. and Sakaki,Y.

Direct Submission

Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical

and Chemical Research (RIKEN), Genomic Sciences Center (GSC);

1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

(E-mail:chimbges@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,

Tel:81-45-503-9111, Fax:81-45-503-9170)

Clones are derived from the chimpanzee BAC library RPCI-43 This BAC

end was generated during the R&D process and may have higher chance

of clone tracking errors.

PRIMERS

Sequencing: T7

LIBRARY

Vector : pBACE3.6

R.Site 1 : EcoRI

R.Site 2 : EcoRI.

Location/Qualifiers

1..716

/organism="Pan troglodytes"

/db\_xref="taxon:9598"

/clone="RP43-030B04.T7"

/sex="male"

/cell\_type="lymphocytes"

/clone\_lib="RPCI-43 Chimpanzee Male BAC Library"

BASE COUNT 117 a 246 c 84 g 268 t

ORIGIN

1 others

Query Match 6.7%; Score 24; DB 12; Length 716;

Best Local Similarity 100.0%; Pred. No. 0.44;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 ggaggaggaggaggaggaggagg 349  
|||||

Db 157 GGAGGAGGAGGAGGAGGAGGAAG 134

## RESULT 7

AI394181  
LOCUS tg67a04.xl Soares\_NHMPu\_S1 Homo sapiens cDNA clone IMAGE:2113806  
DEFINITION 3', mRNA sequence.

ACCESSION AI394181  
VERSION AI394181.1 GI:4223728

KEYWORDS EST.  
SOURCE human.

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

## REFERENCE

1 (bases 1 to 401)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index

## JOURNAL

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)

This clone is available royalty-free through LNL; contact the  
IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.

Insert Length: 480 Std Error: 0.00

Seq primer: -40UP from Gibco.

## FEATURES

Location/Qualifiers

1..401

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:2113806"

/clone\_lib="Soares\_NHMPu\_S1"

/tissue\_type="Pooled human melanocyte, fetal heart, and

pregnant uterus"

/lab\_host="DH10B"

/note="Organ: mixed (see below); Vector: pT73b-Pac  
(Pharmacia) with a modified polylinker; Site\_1: Not I;  
Site\_2: Eco RI; Equal amounts of plasmid DNA from three  
normalized libraries (melanocyte 2N6HM, pregnant uterus  
NbHPU, and fetal heart NbH19W) were mixed, and ss circles  
were made in vitro. Following HAP purification, this DNA  
was used as tracer in a subtractive hybridization  
reaction. The driver was PCR-amplified cDNAs from pools of  
5,000 clones made from the same 3 libraries. The pools  
consisted of I.M.A.G.E. clones 260232-265223,  
340488-345479, and 484488-489479."

BASE COUNT 140 a 61 c 92 g 108 t

## ORIGIN

Query Match 6.4%; Score 23; DB 9; Length 401;

Best Local Similarity 100.0%; Pred. No. 1.3;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 327 gaggaggaggaggaggaggagg 349  
|||||

Db 358 GAGGAGGAGGAGGAGGAGGAAG 380

## RESULT 8

B20638/c B20638 827 bp DNA linear GSS 16-SEP-1997  
LOCUS F5016-T7 IGF Arabidopsis thaliana genomic clone F5016, DNA  
DEFINITION sequence.

ACCESSION B20638

VERSION B20638.1 GI:2395692

KEYWORDS GSS.

SOURCE thale cress.

ORGANISM Arabidopsis thaliana

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## BAC End Sequences at ATGC

## Unpublished (1997)

## Other GSSs: F15016-Sp6

## Contact: Ecker J.

## Arabidopsis Thaliana Genome Center

## University of Pennsylvania

## Dept. of Biology, University of Pennsylvania, Philadelphia, PA

## 19104

## Tel: 215-898-9384

## Fax: 215-898-8780

Email: [jecker@atgenome.bio.upenn.edu](mailto:jecker@atgenome.bio.upenn.edu)

## Seq primer: T7

## Class: BAC ends

## High quality sequence start: 175

## High quality sequence stop: 226.

## FEATURES

## Location/Qualifiers

## 1..827

## /organism="Arabidopsis thaliana"

## /strain="Columbia"

## /db\_xref="taxon:3702"

## /clone="F15016"

## /clone\_lib="IGF"

## /sex="hermaphrodite"

## /note="Vector: BelOBA2II; Site\_1: EcoRI; Site\_2: EcoRI;

## Produced by Thomas Altmann"

## BASE COUNT 197 a 233 c 90 g 284 t

## ORIGIN

## Query Match 6.4%; Score 23; DB 12; Length 827;

## Best Local Similarity 100.0%; Pred. No. 1.4;

## Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 327 gaggaggaggaggaggaggagg 349  
|||||

## Db 791 GAGGAGGAGGAGGAGGAGGAAG 769

## RESULT 9

## BE050974/c

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

## Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC

## clade; Panicoideae; Andropogoneae; Zea.

## 1 (bases 1 to 282)

## O'Shaughnessy, A.L., Habermann, K., de la Bastide, M., Huang, E.N.,

## Nascimben, L.O., Schutz, K., Matero, A., Swaby, I., See, L.-H., Preston

## , R.R., Rodriguez, M.A., Shah, R.S., Shekher, M., Spiegel, L.A., Vil

## , M.D., Dedhia, N.N. and McCombie, W.R.

## Expressed sequence tags from Zea mays (maize)

## Unpublished (2000)

## Contact: W. Richard McCombie

## Lita Annenberg Hazen Genome Sequencing Center

## Cold Spring Harbor Laboratory

## PO Box 100, Cold Spring Harbor, NY 11724, USA

## Tel: 516 367 8884

## Fax: 516 367 8874

Email: [mccombie@cshl.org](mailto:mccombie@cshl.org)

## Plate: za71 row: c column: 12

## Seq primer: -40M13RevUniv

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 827)  
Feng, J., Dewar, K., Buehler, E., Kim, C., Li, Y., Shinn, P., Sun, H. and  
Ecker, J.

## TITLE

## BAC End Sequences at ATGC

## Unpublished (1997)

## Other GSSs: F15016-Sp6

## Contact: Ecker J.

## Arabidopsis Thaliana Genome Center

## University of Pennsylvania

## Dept. of Biology, University of Pennsylvania, Philadelphia, PA

## 19104

## Tel: 215-898-9384

## Fax: 215-898-8780

Email: [jecker@atgenome.bio.upenn.edu](mailto:jecker@atgenome.bio.upenn.edu)

## Seq primer: T7

## Class: BAC ends

## High quality sequence start: 175

## High quality sequence stop: 226.

```
FEATURES
source
High quality sequence stop: 282.
Location/Qualifiers
1..282
/organism="Zea mays"
/db_xref="taxon:4577"
/clone="za71c12"
/clone_lib="Maize Glume cDNAs Library"
/note="Vector: lambda Zap II (Stratagene); Site_1: XhoI;
Site_2: EcoRI; Resistance: Ampicillin Autoexcision:
pBluescript SK (+/-) Titer: 7 x 10e-9 pfu/mL (as of
9/28/94)"
BASE COUNT      36 a 123 c 75 g 48 t
ORIGIN
Query Match      6.2%; Score 22; DB 9; Length 282;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 329 ggaggaggaggaggaggaggc 350
|||||
Db 50 GGAGGGAGGGAGGAGGAAGC 29

RESULT 10
BH293527
LOCUS
DEFINITION
CH230-30L8.TV CHORI-230 Segment 1 Rattus norvegicus genomic clone
CH230-30L8, DNA sequence.
ACCESSION
BH293527
VERSION
BH293527.1 GI:17205935
KEYWORDS
GSS.
SOURCE
Rattus norvegicus
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 309)
Zhao,S., Shetty,J., Shatsman,S., Tsegaye,G., Geer,K., Shvartsbeyn
,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D., Riggs,F., de
Jong,P. and Fraser,C.M.
Rat BAC End Sequences from Library CHORI-230 EcoRI segment
Unpublished (1999)
Other_GSS: CH230-30L8.TJ
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pjejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or ering_information.htm). BAC end
page: http://www.tigr.org/tcdr/bac_ends/rat/bac_end_intro.html
Plate: 30 row: L column: 8
Seq primer: T7
Class: BAC ends.
Location/Qualifiers
1..309
/organism="Rattus norvegicus"
/strain="BN/SsNHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-30L8"
/clone_lib="CHORI-230 Segment 1"
/sex="Female"
/cell_type="Brain"
/note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI;
CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by
Pieter de Jong"
BASE COUNT      114 a 36 c 86 g 73 t

FEATURES
source
Query Match      6.2%; Score 22; DB 10; Length 371;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 328 aggaggaggaggaggaggaggc 349
|||||
Db 166 AGGAGGGAGGGAGGAGGAAG 145

RESULT 12
R11136
LOCUS
DEFINITION
Yf39b09.r1 Soares fetal liver spleen lNFLS Homo sapiens cDNA clone
ORIGIN
Query Match      6.2%; Score 22; DB 12; Length 309;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 328 aggaggaggaggaggaggaggc 349
|||||
Db 278 AGGAGGGAGGGAGGAGGAAG 299

RESULT 11
BF592428/c
LOCUS
DEFINITION
7156a03.x1 NCI_CGAP_Br16 Homo sapiens cDNA clone IMAGE:338668 3'
similar to contains element TARI TARI repetitive element ;, mRNA
sequence.
ACCESSION
BF592428
VERSION
BF592428.1 GI:11684752
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 371)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: Dennis Sgroi, M.D., Kristina Cole, M.D., Ph.D.
student, Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: David B. Krizman, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: -40RP from Gibco.
Location/Qualifiers
1..371
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:338668"
/clone_lib="NCI_CGAP_Br16"
/sex="female"
/tissue_type="lobullar carcinoma in situ"
/lab_host="DH10B"
/lab_stage="adult"
/note="Organ: breast; Vector: pAMP1; mRNA made from breast
carcinoma tissue, cDNA made by oligo-dT priming.
Directionally cloned. Size-selected on agarose gel,
average insert size 400 bp. Primary library,
non-amplified."
BASE COUNT      20 a 184 c 20 g 147 t
ORIGIN
Query Match      6.2%; Score 22; DB 10; Length 371;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 328 aggaggaggaggaggaggaggc 349
|||||
Db 166 AGGAGGGAGGGAGGAGGAAG 145

RESULT 12
R11136
LOCUS
DEFINITION
Yf39b09.r1 Soares fetal liver spleen lNFLS Homo sapiens cDNA clone
```



JOURNAL  
COMMENT

Unpublished (2000)  
Contact: W. Richard McCombie  
Lita Annenberg Hazen Genome Sequencing Center  
Cold Spring Harbor Laboratory  
PO Box 100, Cold Spring Harbor, NY 11724, USA  
Tel: 516 367 8884  
Fax: 516 367 8874  
Email: mcombie@cshl.org

Plate: za71 row: c column: 12  
Seq primer: -40M3ForUniv  
High quality sequence stop: 478.

FEATURES  
source

Location/Qualifiers  
1. .478  
/organism="Zea mays"  
/db\_xref="taxon:4577"  
/clone="za71c12"  
/clone\_lib="Maize Glume cDNAs Library"  
/note="Vector: Lambda zap II (Stratagene); Site\_1: XhoI;  
Site\_2: EcoRI; Resistance: Ampicillin Autoexcision;  
pBluescript SK (+/-) Titer: 7 x 10e-9 pfu/mL (as of  
9/28/94)."

BASE COUNT 89 a 170 c 127 g 92 t  
ORIGIN

Query Match 6.2%; Score 22; DB 9; Length 478;  
Best Local Similarity 100.0%; Pred. No. 4.1; 0; Indels 0; Gaps 0;  
Matches 22; Conservative 0; Mismatches 0

QY 329 ggaggaggaggaggaggaggc 350  
|||||  
Db 44 GGAGGGAGGGAGGAGGAGC 23

RESULT 15  
AZ654841/c

LOCUS 483 bp DNA linear GSS 14-DEC-2000  
DEFINITION M0529P18F Mouse 10kb plasmid UUGCLM library Mus musculus genomic  
clone UUGCLM0529P18 F, DNA sequence.

ACCESSION AZ654841  
VERSION 1 GI:11791987  
KEYWORDS GSS.

## SOURCE

house mouse  
Mus musculus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
1 (bases 1 to 483)  
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,  
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,  
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.  
and Wright, D. Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

JOURNAL Contact: Robert B. Weiss  
COMMENT University of Utah Genome Center  
University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0529 row: p column: 18

Seq primer: CGTGTAAACGACGCCAGT

Class: plasmid ends

High quality sequence stop: 483.

Location/Qualifiers

1. .483

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="UUGCLM0529P18"

FEATURES  
source

Location/Qualifiers  
1. .515

FEATURES  
source

Location/Qualifiers  
1. .515

/clone\_lib="Mouse 10kb plasmid UUGCLM library"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/note="Vector: PWD42nv; Purified genomic DNA from M.  
musculus C57BL/6J (male) was obtained from the Jackson  
Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA  
was hydrodynamically sheared by repeated passage through a  
0.005 inch orifice at constant velocity. The sheared DNA  
was blunt end-repaired with T4 DNA polymerase and T4  
polynucleotide kinase. Adaptor oligonucleotides were  
ligated to the blunt ends in high molar excess. The  
adapted DNA was purified and size-selected for a 9.5 to  
10.5 kb range using preparative agarose gel  
electrophoresis. Vector DNA was prepared from a derivative  
of PWD42 (gi14732114|gb|AF129072.1), a copy-number  
inducible derivative of plasmid R1. The vector was ligated  
with adaptors complementary to the insert adaptors and  
purified. The sheared, adapted mouse DNA was annealed to  
adapted vector DNA, and transformed into  
chemically-competent E. coli XL10-Gold (Stratagene) cells  
and selected for ampicillin resistance."

BASE COUNT 107 a 140 c 79 g 157 t  
ORIGIN

Query Match 6.2%; Score 22; DB 12; Length 483;  
Best Local Similarity 100.0%; Pred. No. 4.1; 0; Indels 0; Gaps 0;  
Matches 22; Conservative 0; Mismatches 0

QY 326 ggaggaggaggaggaggaggc 347  
|||||  
Db 340 GGAGGGAGGGAGGAGGAGC 319

RESULT 16  
BH337839/c

LOCUS 515 bp DNA linear GSS 03-DEC-2001  
DEFINITION BH337839 TV CHORI-230 Segment 1 Rattus norvegicus genomic clone  
CH230-194J9, DNA sequence.

ACCESSION BH337839  
VERSION 1 GI:17268573  
KEYWORDS GSS.

## SOURCE

Norway rat.  
Rattus norvegicus

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

REFERENCE 1 (bases 1 to 515)  
AUTHORS Zhao, S., Shetty, J., Shatsman, S., Tsegaye, G., Geer, K., Shvartsbeyn,  
A., Gebregorgis, E., Overton, L., Russell, D., Chen, D., Riggs, F., de  
Jong, P. and Fraser, C. M.

Rat BAC End Sequences from Library CHORI-230 EcoRI segment

Unpublished (1999)

Other\_GSSs: CH230-194J9.TJ

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the rat BAC library CHORI-230

(http://www.chori.org/bacpac/rat230.htm). For BAC library

availability, please contact Pieter de Jong (pjejong@mail.cho.org).

Clones may be purchased from BACPAC Resources

(http://www.chori.org/bacpac/or ering\_information.htm). BAC end

page: http://www.tigr.org/tdb/bac\_ends/rat/bac\_end\_intro.html

Plate: 194 row: J column: 9

Seq primer: T7

Class: BAC ends.

Location/Qualifiers

1. .515

```

/organism="Rattus norvegicus"
/strain="BN/SSNHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-194J9"
/clone_lib="CHORI-230 Segment 1"
/sex="Female"
/cell_type="Brain"
/notes="Vector: pTARAC2.1; Site_1: EcoRI; Site_2: EcoRI;
CHORI-230 Rat (BN/SSNHsd/MCW) BAC library produced by
Pieter de Jong"
BASE COUNT      152 a 154 c 90 g 119 t
ORIGIN

Query Match      6.2%; Score 22; DB 12; Length 515;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 gagagaggaggaggaggagga 347
|||||
DB 243 GGAGGAGGAGGAGGAGGA 222

RESULT 17
LOCUS      BM310687/c
DEFINITION      ig47b12.y1 HR85 islet Homo sapiens cDNA 5', mRNA sequence.
ACCESSION      BM310687
VERSION      BM310687.1 GI:18045003
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 523)
Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
Lemishka, I., Seearce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,
Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas
, M., Gibbons, M., McCann, R., Cole, R., Tsagarelshvili, R., Williams, T.,
Jackson, Y. and Bowers, Y.
Endocrine Pancreas Consortium
Other ESTs: ig47b12.x1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)
Seq primer: -408P from Gibco
High quality sequence stop: 488.
FEATURES
source
1..523
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HR85 islet"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/notes="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
Size-selected on agarose gel. Average insert size -1kb. 5'
XhoI site was destroyed after directional cloning.
Amplified once. Contact information: Hiroshi Inoue, MD,
Metabolism Div. (Alan Permutt Lab), Washington University
School of Medicine, Box 8127, 660 South Euclid Ave., St.
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:

/organism="Rattus norvegicus"
/strain="BN/SSNHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-194J9"
/clone_lib="CHORI-230 Segment 1"
/sex="Female"
/cell_type="Brain"
/notes="Vector: pTARAC2.1; Site_1: EcoRI; Site_2: EcoRI;
CHORI-230 Rat (BN/SSNHsd/MCW) BAC library produced by
Pieter de Jong"
BASE COUNT      152 a 154 c 90 g 119 t
ORIGIN

Query Match      6.2%; Score 22; DB 12; Length 515;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 gagagaggaggaggaggagga 347
|||||
DB 243 GGAGGAGGAGGAGGAGGA 222

RESULT 17
LOCUS      BM310687/c
DEFINITION      ig47b12.y1 HR85 islet Homo sapiens cDNA 5', mRNA sequence.
ACCESSION      BM310687
VERSION      BM310687.1 GI:18045003
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 523)
Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,
Lemishka, I., Seearce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,
Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas
, M., Gibbons, M., McCann, R., Cole, R., Tsagarelshvili, R., Williams, T.,
Jackson, Y. and Bowers, Y.
Endocrine Pancreas Consortium
Other ESTs: ig47b12.x1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)
Seq primer: -408P from Gibco
High quality sequence stop: 488.
FEATURES
source
1..523
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HR85 islet"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/notes="Organ: Pancreas; Vector: pBluescript SK(-); Site_1:
NotI; Site_2: XhoI; cDNA made by oligo-dT priming.
Size-selected on agarose gel. Average insert size -1kb. 5'
XhoI site was destroyed after directional cloning.
Amplified once. Contact information: Hiroshi Inoue, MD,
Metabolism Div. (Alan Permutt Lab), Washington University
School of Medicine, Box 8127, 660 South Euclid Ave., St.
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGCLM0513L22"
/clone_lib="Mouse 10kb plasmid UUGCLM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PWD42hiv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (gii4732114|gbiAF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into

314-362-1916, Fax: 314-747-2692."
109 a 127 c 88 g 199 t
BASE COUNT      109 a 127 c 88 g 199 t
ORIGIN

Query Match      6.2%; Score 22; DB 10; Length 523;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 327 gagagaggaggaggaggagga 348
|||||
DB 291 GAGGAGGAGGAGGAGGA 270

RESULT 18
LOCUS      AZ647100/c
DEFINITION      1M0513L22F Mouse 10kb plasmid UUGCLM library Mus musculus genomic
clone UUGCLM0513L22 F, DNA sequence.
ACCESSION      AZ647100
VERSION      AZ647100.1 GI:11778230
KEYWORDS      GSS.
SOURCE      house mouse.
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 525)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly
, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
and Wright, D., Weiss, R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: dduenne@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0513 row: L column: 22
Seq primer: CTTGTAAACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 525.
FEATURES
Location/Qualifiers
1..525
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGCLM0513L22"
/clone_lib="Mouse 10kb plasmid UUGCLM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PWD42hiv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (gii4732114|gbiAF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into

```

chemically-competent *E. coli* XL10-Gold (Stratagene) cells  
and selected for ampicillin resistance."

BASE COUNT  
ORIGIN

95 a 156 c 91 g 183 t

Query Match

Best Local Similarity 6.2%; Score 22; DB 12; Length 525;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 328 agagaggaggaggaggaggag 349

Db 347 AGGAGGAGGAGGAGGAGGAG 326

RESULT 19

LOCUS

DEFINITION A0415718 544 bp DNA linear GSS 23-MAR-1999

RPCI-11-204G24-TV RPCI-11 Homo sapiens genomic clone RPCI-11-204G24

, DNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Other\_GSSs: RPCI-11-204G24.TJ

Contact: Shaying Zhao, William Nierman, Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850

Tel: 301 838 0200

Fax: 301 838 0208

Email: hbe@tigr.org

Clones are derived from the human BAC library RPCI-11. For BAC

library availability, please contact Pieter de Jong

(pieter@dejong.med.buffalo.edu). Clones may be purchased from

BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from

Research Genet cs (info@resgen.com). BAC end search page:

http://www.tigr.org/tcdb/humgen/bac\_end\_search/bac\_end\_search.html.

Seq primer: T7

Class: BAC ends.

FEATURES

source

Location/Qualifiers

1..544

/organism="Homo sapiens"

/db\_xref="GDB:7578119"

/db\_xref="taxon:9606"

/clone="RPCI-11-204G24"

/clone\_lib="RPCI-11"

/sex="Male"

/cell\_type="Lymphocytes"

/note="Vector: pBACe3.6; Site\_1: EcoRI; Site\_2: EcoRI;

RPCI11 Human Male BAC Library"

90 a 162 c 78 g 214 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 6.2%; Score 22; DB 12; Length 544;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 328 agagaggaggaggaggaggag 349

Db 182 AGGAGGAGGAGGAGGAGGAG 161

RESULT 20

LOCUS

DEFINITION

AV938157 K. Sato unpublished cDNA library, strain H602 adult,

heading stage top three leaves Hordeum vulgare subsp. spontaneum

cDNA clone bahl6g20 5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

Location/Qualifiers

1..552

/organism="Hordeum vulgare subsp. spontaneum"

/strain="H602"

/db\_xref="taxon:77009"

/clone="bahl6g20"

/clone\_lib="K. Sato unpublished cDNA library, strain H602

adult, heading stage top three leaves"

/tissue\_type="top three leaves"

/dev\_stage="adult, heading stage"

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 6.2%; Score 22; DB 9; Length 552;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 326 ggaggaggaggaggaggaggag 347

Db 37 GGAGGAGGAGGAGGAGGAGGA 16

RESULT 21

LOCUS

DEFINITION

AQ078910 569 bp DNA linear GSS 20-AUG-1998

CIT-HSP-2367J11.TR CIT-HSP Homo sapiens genomic clone 2367J11, DNA

sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Use of a random human BAC End Sequence Database for Sequence-Ready

Map Building

Unpublished (1998)

Other\_GSSs: CIT-HSP-2367J11.TF

Contact: Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

AV938157 552 bp mRNA linear EST 18-JAN-2002

AV938157 K. Sato unpublished cDNA library, strain H602 adult,

heading stage top three leaves Hordeum vulgare subsp. spontaneum

cDNA clone bahl6g20 5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Tadasu Shin-i

Center For Genetic Resource Information

National Institute of Genetics

1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856

Fax: 81-559-81-6855

Email: tshini@genes.nig.ac.jp.

Location/Qualifiers

1..552

/organism="Hordeum vulgare subsp. spontaneum"

/strain="H602"

/db\_xref="taxon:77009"

/clone="bahl6g20"

/clone\_lib="K. Sato unpublished cDNA library, strain H602

adult, heading stage top three leaves"

/tissue\_type="top three leaves"

/dev\_stage="adult, heading stage"

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 6.2%; Score 22; DB 9; Length 552;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 326 ggaggaggaggaggaggaggag 347

Db 37 GGAGGAGGAGGAGGAGGAGGA 16

RESULT 21

LOCUS

DEFINITION

AQ078910 569 bp DNA linear GSS 20-AUG-1998

CIT-HSP-2367J11.TR CIT-HSP Homo sapiens genomic clone 2367J11, DNA

sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Use of a random human BAC End Sequence Database for Sequence-Ready

Map Building

Unpublished (1998)

Other\_GSSs: CIT-HSP-2367J11.TF

Contact: Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208  
 Email: madams@tigr.org  
 Clones are available from Research Genetics (info@resgen.com). BAC  
 end search page:  
[http://www.tigr.org/tdb/hungen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html).  
 Seq primer: M13 Reverse  
 Class: BAC ends.

## FEATURES

source Location/Qualifiers  
 1. 569  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="2367J11"  
 /clone\_lib="CIT-HSP"  
 /sex="Male"  
 /cell\_type="Sperm"  
 /note="Vector: pbelOBAC11; Site\_1: HindIII; Site\_2:  
 HindIII"

BASE COUNT 68 a 189 c 57 g 254 t 1 others  
 ORIGIN

Query Match 6.2%; Score 22; DB 12; Length 569;  
 Best Local Similarity 100.0%; Pred. No. 4.2;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 328 aggaggaggaggaggaggagg 349  
 |||||  
 Db 454 AGGAGGAGGAGGAGGAGGAG 433

## RESULT 22

AZ378099/c 572 bp DNA linear GSS 02-OCT-2000  
 LOCUS  
 DEFINITION  
 clone UGCLM0132P08 R, DNA sequence.

ACCESSION AZ378099  
 VERSION  
 KEYWORDS  
 SOURCE

ORGANISM  
 house mouse.  
 Mus musculus  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 1 (bases 1 to 572)  
 Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly  
 ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.  
 and Wright,D., Weiss,R.  
 Mouse whole genome scaffolding with paired end reads from 10kb  
 plasmid inserts

TITLE  
 JOURNAL  
 COMMENT  
 Unpublished (2000)  
 Contact: Robert B. Weiss  
 University of Utah Genome Center  
 University of Utah  
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
 84112, USA  
 Tel: 801 585 5606  
 Fax: 801 585 7177  
 Email: ddunn@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Plate: 0132 row: P column: 08  
 Seq primer: CACACAGGAACAGCTATGACC  
 Class: plasmid ends  
 High quality sequence stop: 572.

## FEATURES

source Location/Qualifiers  
 1. 572  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UGCLM0132P08"  
 /clone\_lib="Mouse 10kb plasmid UGCLM library"  
 /sex="Male"  
 /lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
 /note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson  
 Laboratory Mouse DNA Resource  
 (<http://www.jax.org/resources/documents/dnares/>). The DNA  
 was hydrodynamically sheared by repeated passage through a  
 0.005 inch orifice at constant velocity. The sheared DNA  
 was blunt end-repaired with T4 DNA polymerase and T4  
 polynucleotide kinase. Adaptor oligonucleotides were  
 ligated to the blunt ends in high molar excess. The  
 adaptor DNA was purified and size-selected for a 9.5 to  
 10.5 kb range using preparative agarose gel  
 electrophoresis. Vector DNA was prepared from a derivative  
 of pWD42 (gil4732114|gb|AF129072.1), a copy-number  
 inducible derivative of plasmid R1. The vector was ligated  
 with adaptors complementary to the insert adaptors and  
 purified. The sheared, adaptor mouse DNA was annealed to  
 adaptor vector DNA, and transformed into  
 chemically-competent E. coli XL10-Gold (Stratagene) cells  
 and selected for ampicillin resistance."

BASE COUNT 130 a 137 c 118 g 187 t  
 ORIGIN

Query Match 6.2%; Score 22; DB 12; Length 572;  
 Best Local Similarity 100.0%; Pred. No. 4.2;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 328 aggaggaggaggaggaggagg 349  
 |||||  
 Db 231 AGGAGGAGGAGGAGGAGGAG 210

## RESULT 23

AZ828464/c 591 bp DNA linear GSS 20-FEB-2001  
 LOCUS  
 DEFINITION  
 clone UGCG2M0105N23 F, DNA sequence.

ACCESSION AZ828464  
 VERSION  
 KEYWORDS  
 SOURCE

ORGANISM  
 house mouse.  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 1 (bases 1 to 591)  
 Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly  
 ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.  
 and Wright,D., Weiss,R.  
 Mouse whole genome scaffolding with paired end reads from 10kb  
 plasmid inserts

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT  
 Unpublished (2000)  
 Contact: Robert B. Weiss  
 University of Utah Genome Center  
 University of Utah  
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
 84112, USA  
 Tel: 801 585 5606  
 Fax: 801 585 7177  
 Email: ddunn@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Plate: 0103 row: N column: 23  
 Seq primer: CGTTGTAAACGACGCCAGT  
 Class: plasmid ends  
 High quality sequence stop: 591.

## FEATURES

source Location/Qualifiers  
 1. 591  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UGCG2M0105N23"  
 /clone\_lib="Mouse 10kb plasmid UGCLM library"  
 /sex="Male"  
 /lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"



/note="Vector: pWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 117 a 189 c 110 g 175 t  
ORIGIN

Query Match 6.2%; Score 22; DB 12; Length 591;  
Best Local Similarity 100.0%; Pred. No. 4.2;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 ggaaggaggaggaggaggaggc 350  
|||||  
Db 130 GGAGGGAGGGAGGAGGAGGAGC 109

RESULT 24  
AQ924975/c  
LOCUS  
DEFINITION AQ924975 615 bp DNA linear GSS 21-DEC-1999  
, DNA sequence.  
ACCESSION AQ924975.1 GI:6613978  
VERSION  
KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
REFERENCE 1 (bases 1 to 615)  
AUTHORS Zhao,S., Niernan,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret ,B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P. and Fraser,C.M.  
TITLE Mouse BAC End Sequences from Library RPCI-23  
JOURNAL Unpublished (1999)  
COMMENT Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: szhao@tigr.org  
Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>) or from Resea ch Genetics (info@resgen.com). BAC end page: [http://www.tigr.org/tdb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html)  
Plate: 277 row: E column: 11  
Seq primer: T7  
Class: BAC ends.

FEATURES  
source  
1. .615  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="RPCI-23-277E11"  
/clone\_lib="RPCI-23"

/sex="female"  
/lab\_host="DH10B"  
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site\_1: EcoRI; Site\_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACe3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 144 a 158 c 101 g 212 t  
ORIGIN

Query Match 6.2%; Score 22; DB 12; Length 615;  
Best Local Similarity 100.0%; Pred. No. 4.2;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 328 aggaaggaggaggaggaggaggc 349  
|||||  
Db 506 AGGAGGGAGGAGGAGGAGGAG 485

RESULT 25  
B70930/c  
LOCUS  
DEFINITION B70930 623 bp DNA linear GSS 21-JUN-1998  
CIT-HSP-2063G10.TF CIT-HSP Homo sapiens genomic clone 2063G10, DNA sequence.

ACCESSION B70930  
VERSION B70930.1 GI:2710154  
KEYWORDS GSS.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 623)  
AUTHORS Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Golden ,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and Venter,J.C.  
TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map Building  
JOURNAL Unpublished (1997)  
COMMENT Other\_GSSs: CIT-HSP-2063G10.TR  
Contact: Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: mdadams@tigr.org  
Clones are available from Research Genetics (info@resgen.com). BAC end search page: [http://www.tigr.org/tdb/hungen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html)  
Seq primer: M13-21  
Class: BAC ends.

FEATURES  
Location/Qualifiers  
1. .623  
/organism="Homo sapiens"  
/db\_xref="GDB:706153"  
/db\_xref="taxon:9606"  
/clone="2063G10"  
/clone\_lib="CIT-HSP"  
/sex="Male"  
/cell\_type="Sperm"  
/note="Vector: pBelOBAC11; Site\_1: HindIII; Site\_2: HindIII"

BASE COUNT 74 a 209 c 72 g 268 t  
ORIGIN

Query Match 6.2%; Score 22; DB 12; Length 623;  
Best Local Similarity 100.0%; Pred. No. 4.2;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 328 agggaggaggaggaggaggaggagg 349
Db 491 AGGAGGAGGAGGAGGAGGAGGAG 470

RESULT 26
BF984535
LOCUS 602307723F1 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:4399276 5',
DEFINITION mRNA sequence.
ACCESSION BF984535
VERSION BF984535.1 GI:12387347
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 673)
AUTHORS NIH-MGC http://mhc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CNA Library Preparation: Life Technologies, Inc.
CNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLAM10102 row: m column: 05
High quality sequence stop: 671.
FEATURES
    source
        1..673
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="IMAGE:4399276"
        /clone_lib="NIH_MGC_88"
        /tissue_type="duodenal adenocarcinoma, cell line"
        /lab_host="DH10B (phage-resistant)"
        /note="Organ: small intestine; Vector: pCMV-SPORT6;
        Site 1: NotI; Site 2: SalI; Cloned unidirectionally;
        oligo-dT primed. Average insert size 1.767 kb. Library
        enriched for full-length clones and constructed by Life
        Technologies. Note: this is a NIH_MGC Library."
BASE COUNT 150 a 184 c 204 g 135 t
ORIGIN

Query Match 6.2%; Score 22; DB 10; Length 573;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 gggaggaggaggaggaggaggagg 347
Db 33 GGAGGAGGAGGAGGAGGAGGAGGA 54

RESULT 27
AZ742461/c
LOCUS AZ742461
DEFINITION RPCI-24-74G13.TJ RPCI-24 Mus musculus genomic clone RPCI-24-74G13,
DNA sequence.
ACCESSION AZ742461
VERSION AZ742461.1 GI:12521331
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 713)
AUTHORS Zhao, S., Nierman, W., Malek, J., Shatsman, S., Akinret, B.,
Tsegaye, G., Geer, K., Krol, M., Shvartsbeyn, A., Gebregeorgis, E.,

Russell, D., de Jong, P. and Fraser, C.M.
Mouse BAC End Sequences from Library RPCI-24
Unpublished (1999)
Other GSSs: RPCI-24-74G13.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-24. For BAC
library availability, please contact Pieter de Jong
(pdejong@mail.cho.org). Clones may be purchased from BACPAC
resources (http://www.choi.org/bacpac/orderingframe.htm). BAC end
page: http://ww.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 74 row: G column: 13
Seq primer: SP6
Class: BAC ends.
FEATURES
    source
        1..713
        /organism="Mus musculus"
        /strain="C57BL/6J"
        /db_xref="taxon:10090"
        /clone="RPCI-24-74G13"
        /clone_lib="RPCI-24"
        /sex="Male"
        /cell_type="Spleen/Brain"
        /note="Vector: pTARBAC1; Site 1: BamHI; Site 2: BamHI;
        RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
        library was cloned in the pTARBAC1 cloning vector at the
        BamHI sites using MboI partially digested male C57BL/6J
        DNA."
BASE COUNT 126 a 225 c 114 g 248 t
ORIGIN

Query Match 6.2%; Score 22; DB 12; Length 713;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 328 agggaggaggaggaggaggaggagg 349
Db 446 AGGAGGAGGAGGAGGAGGAGGAG 425

RESULT 28
AQ983682/c
LOCUS AQ983682
DEFINITION RPCI-23-323J22.TJ RPCI-23 Mus musculus genomic clone RPCI-23-323J22
, DNA sequence.
ACCESSION AQ983682
VERSION AQ983682.1 GI:6816887
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 754)
AUTHORS Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S., Akinret,
B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P.,
and Fraser, C.M.
Mouse BAC End Sequences from Library RPCI-23
Unpublished (1999)
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC
library availability, please contact Pieter de Jong

```

(pieter@dejong.med.buffalo.edu). Clones may be purchased from BACpac Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>) or from Resea ch Genetics ([inforesgen.com](http://inforesgen.com)). BAC end page: [http://www.tigr.org/tldb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html)  
 Plate: 323 row: J' column: 22  
 Seq primer: SP6  
 Class: BAC ends.

#### FEATURES

Location/Qualifiers  
 1..754  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="RPCI-23-323J22"  
 /clone\_lib="RPCI-23"  
 /sex="Female"  
 /lab\_host="DH10B"

/note="Organ: Kidney/Brain; Vector: pBACE3.6; Site\_1: EcoRI; Site\_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACE3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."  
 182 a 189 c 131 g 251 t 1 others

#### BASE COUNT

182 a 189 c 131 g 251 t 1 others

Query Match 6.2%; Score 22; DB 12; Length 754;  
 Best Local Similarity 100.0%; Pred. No. 4.3;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 328 agaggaggaggaggaggagg 349  
 Db 510 AGGAGGGAGGGAGGGAAG 489

#### RESULT 29

BE877424 760 bp mRNA linear EST 20-OCT-2000  
 LOCUS 601485462F1 NIH\_MGC\_69 Homo sapiens cDNA clone IMAGE:3888134 5',  
 DEFINITION mRNA sequence.

ACCESSION BE877424  
 VERSION BE877424.1 GI:10326200  
 KEYWORDS EST.  
 SOURCE human.

#### ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 760)

NIH-MGC <http://mgc.nci.nih.gov/>  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)

Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov

Tissue Procurement: DCTD/DMP/Gazdar

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLN9667 row: k column: 15

High quality sequence stop: 678.

Location/Qualifiers

1..760

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:3888134"

/clone\_lib="NIH\_MGC\_69"

/tissue\_type="large cell carcinoma, undifferentiated"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: lung; Vector: pCMV-SPORT6; Site\_1: NotI;  
 Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.

#### FEATURES

Location/Qualifiers  
 1..760

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:3888134"

/clone\_lib="NIH\_MGC\_69"

/tissue\_type="large cell carcinoma, undifferentiated"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: lung; Vector: pCMV-SPORT6; Site\_1: NotI;  
 Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.1 kb. Library constructed by Life Technologies."

BASE COUNT 173 a 207 c 233 g 147 t  
 ORIGIN

Query Match 6.2%; Score 22; DB 10; Length 760;  
 Best Local Similarity 100.0%; Pred. No. 4.3;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 ggaggaggaggaggaggagg 347  
 Db 14 GGAGGAGGGAGGAGGGAAGGA 35

#### RESULT 30

BE617281 907 bp mRNA linear EST 20-OCT-2000  
 LOCUS 601441977F1 NIH\_MGC\_65 Homo sapiens cDNA clone IMAGE:3846235 5',  
 DEFINITION mRNA sequence.

ACCESSION BE617281

VERSION BE617281.1 GI:9888219

KEYWORDS EST.

SOURCE human.

#### ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 907)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLN9558 row: i column: 20

High quality sequence start: 8

High quality sequence stop: 660.

Location/Qualifiers

1..907

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:3846235"

/clone\_lib="NIH\_MGC\_65"

/tissue\_type="adenocarcinoma"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: colon; Vector: pCMV-SPORT6; Site\_1: NotI;  
 Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.8 kb. Library constructed by Life Technologies."

BASE COUNT 197 a 252 c 304 g 154 t  
 ORIGIN

Query Match 6.2%; Score 22; DB 10; Length 907;  
 Best Local Similarity 100.0%; Pred. No. 4.4;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 326 ggaggaggaggaggaggagg 347  
 Db 15 GGAGGAGGGAGGAGGGAAGGA 36

#### RESULT 31

BG823588 926 bp mRNA linear EST 22-MAY-2001  
 LOCUS 602728977F1 NIH\_MGC\_15 Homo sapiens cDNA clone IMAGE:4868443 5',  
 DEFINITION mRNA sequence.

```

ACCESSION      BG823588
VERSION         BG823588.1  GI:14171175
KEYWORDS       EST.
SOURCE         human.
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS        National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE          Unpublished (1999)
JOURNAL        Contact: Robert Strausberg, Ph.D.
COMMENT        Email: cgabs@email.nih.gov
               Tissue Procurement: ATCC
               cDNA Library Preparation: Ling Hong/Rubin Laboratory
               cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
               DNA Sequencing by: NIH Intramural Sequencing Center
               Clone distribution: MGC clone distribution information can be
               found through the I.M.A.G.E. Consortium/LLNL at:
               http://image.llnl.gov
               Plate: LLCMI738 row: i column: 20
               High quality sequence stop: 728.
               Location/Qualifiers
               1..926
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /clone="IMAGE:4868443"
               /clone_lib="NIH_MGC_15"
               /tissue_type="adenocarcinoma cell line"
               /lab_host="DH10B (phage-resistant)"
               /note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2:
               EcoRI; cDNA made by oligo-dT priming. Directionally
               cloned into EcoRI/XhoI sites using the following 5',
               adaptor: GGCACGAG(G). Size-selected >500bp for average
               insert size 1.8kb. Library constructed by Ling Hong in
               the laboratory of Gerald M. Rubin (University of
               California, Berkeley) using ZAP-cDNA synthesis kit
               (Stratagene) and Superscript II RT (Life Technologies)"
               208 a 253 c 298 g 167 t

BASE COUNT     208 a 253 c 298 g 167 t
ORIGIN

Query Match    6.2%; Score 22; DB 10; Length 926;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 326 ggagggaggaggagggaagga 347
|||||
Db 10 GGAGGAGGAGGAGGGAAGGA 31

RESULT 32
BG328313
LOCUS          602427291F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:4546583 5',
DEFINITION    mRNA sequence.
ACCESSION     BG328313
VERSION       BG328313.1  GI:13134660
KEYWORDS      EST.
SOURCE        human.
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS        National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE          Unpublished (1999)
JOURNAL        Contact: Robert Strausberg, Ph.D.
COMMENT        Email: cgabs@email.nih.gov
               Tissue Procurement: ATCC
               cDNA Library Preparation: Ling Hong/Rubin Laboratory
               cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
               DNA Sequencing by: NIH Intramural Sequencing Center

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCMI738 row: f column: 24
High quality sequence stop: 668.
Location/Qualifiers
1..1068
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4546583"
/clone_lib="NIH_MGC_15"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: colon; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5',
adaptor: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)"
261 a 294 c 332 g 181 t

BASE COUNT     261 a 294 c 332 g 181 t
ORIGIN

Query Match    6.2%; Score 22; DB 10; Length 1068;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 326 ggagggaggaggagggaagga 347
|||||
Db 6 GGAGGAGGAGGAGGGAAGGA 27

RESULT 33
A2995124/c
LOCUS          2M0280018R Mouse 10kb plasmid UUGC2M library Mus musculus genomic
DEFINITION    clone UUGC2M0280018 R, DNA sequence.
ACCESSION     A2995124
VERSION       A2995124.1  GI:13866351
KEYWORDS      GSS.
SOURCE        house mouse.
ORGANISM      Mus musculus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE      1 (bases 1 to 43)
AUTHORS        Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
               Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
               ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
               and Wright,D., Weiss,R.
               Mouse whole genome scaffolding with paired end reads from 10kb
               plasmid inserts
               Unpublished (2000)
               Contact: Robert B. Weiss
               University of Utah Genome Center
               Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
               84112, USA
               Tel: 801 585 5606
               Fax: 801 585 7177
               Email: ddunne@genetics.utah.edu
               Insert Length: 10000 Std Error: 0.00
               Plate: 0280 row: 0 column: 18
               Seq primer: CACACAGAAACAGCTATGACC
               Class: plasmid ends
               High quality sequence stop: 43.
               Location/Qualifiers
               1..43
               /organism="Mus musculus"
               /strain="C57BL/6J"
               /db_xref="taxon:10090"
               /clone="UUGC2M0280018"

```

```

/clone_lib="Mouse 10kb plasmid UUGC2M library"
/sex="Female"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gii47321141gb)AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      0 a      27 c      2 g      14 t
ORIGIN

```

Query Match 5.9%; Score 21; DB 12; Length 43;

Best Local Similarity 100.0%; Pred. No. 10; Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 ggaggaggaggaggaggagg 349

Db 34 GGAGGAGGGGAGGAGGAG 14

RESULT 34

AZ351915

LOCUS

DEFINITION 1M0090K10F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0090K10 F, DNA sequence.

ACCESSION AZ351915

VERSION AZ351915.1 GI:10431152

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM

Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 51)

REFERENCE

AUTHORS

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmood,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

Unpublished (2000)

JOURNAL

COMMENT

Contact: Robert B. Weiss

University of Utah Genome Center

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0090 row: K column: 10

Seq primer: CGTTGTAACGACGCGCCAGT

Class: plasmid ends

High quality sequence stop: 51.

Location/Qualifiers

1..51

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

FEATURES

source

```

/clone_lib="UUGC1M0090K10"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gii47321141gb)AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      16 a      0 c      35 g      0 t
ORIGIN

```

Query Match 5.9%; Score 21; DB 12; Length 51;

Best Local Similarity 100.0%; Pred. No. 10; Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 ggaggaggaggaggaggagg 349

Db 19 GGAGGAGGGGAGGAGGAG 39

RESULT 35

AZ500262

LOCUS

DEFINITION 1M0338H04R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0338H04 R, DNA sequence.

ACCESSION AZ500262

VERSION AZ500262.1 GI:10679897

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM

Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 62)

REFERENCE

AUTHORS

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmood,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

Unpublished (2000)

JOURNAL

COMMENT

Contact: Robert B. Weiss

University of Utah Genome Center

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0338 row: H column: 04

Seq primer: CACACAGGAACAGCATGACC

Class: plasmid ends

High quality sequence stop: 62.

Location/Qualifiers

1..62

/organism="Mus musculus"

/strain="C57BL/6J"

```

/db_xref="taxon:10090"
/clone="UUGC1M0338H04"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid RL. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      27 a 1 c 32 g 2 t
ORIGIN

```

```

Query Match      5.9%; Score 21; DB 12; Length 62;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 329 ggaggaggaggaggaggaggagg 349
      |||||
Db 26 GGAGGGAGGGAGGAGGAGAG 46

```

```

RESULT 36
AZ411857/c
LOCUS
DEFINITION
  M0185006F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
  clone UUGC1M0185006 F, DNA sequence.
ACCESSION
  AZ411857
VERSION
  AZ411857.1 GI:10535870
KEYWORDS
  GSS.
SOURCE
  house mouse.
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1 (bases 1 to 66)
REFERENCE
  AUTHORS
    Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
    Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
    ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
    and Wright,D., Weiss,R.
  TITLE
    Mouse whole genome scaffolding with paired end reads from 10kb
    plasmid inserts
  JOURNAL
    Unpublished (2000)
  COMMENT
    Contact: Robert B. Weiss
    University of Utah Genome Center
    Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
    84112, USA
    Tel: 801 585 5606
    Fax: 801 585 7177
    Email: ddunn@genetics.utah.edu
    Insert Length: 10000 Std Error: 0.00
    Plate: 0185 row: 0 column: 06
    Seq primer: CGTTGTAACGACGCCAGT
    Class: plasmid ends
    High quality sequence stop: 66.
    Location/Qualifiers
      1..66
      /organism="Mus musculus"

```

```

FEATURES
  source
    1..66
    /organism="Mus musculus"

```

```

/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0185006"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid RL. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      3 a 37 c 2 g 24 t
ORIGIN

```

```

Query Match      5.9%; Score 21; DB 12; Length 66;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 329 ggaggaggaggaggaggaggagg 349
      |||||
Db 57 GGAGGGAGGGAGGAGGAGAG 37

```

```

RESULT 37
AZ838528/c
LOCUS
DEFINITION
  M0134E17F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
  clone UUGC2M0134E17 F, DNA sequence.
ACCESSION
  AZ838528
VERSION
  AZ838528.1 GI:13008436
KEYWORDS
  GSS.
SOURCE
  house mouse.
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1 (bases 1 to 90)
REFERENCE
  AUTHORS
    Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
    Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
    ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
    and Wright,D., Weiss,R.
  TITLE
    Mouse whole genome scaffolding with paired end reads from 10kb
    plasmid inserts
  JOURNAL
    Unpublished (2000)
  COMMENT
    Contact: Robert B. Weiss
    University of Utah Genome Center
    Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
    84112, USA
    Tel: 801 585 5606
    Fax: 801 585 7177
    Email: ddunn@genetics.utah.edu
    Insert Length: 10000 Std Error: 0.00
    Plate: 0134 row: E column: 17
    Seq primer: CGTTGTAACGACGCCAGT
    Class: plasmid ends
    High quality sequence stop: 90.
    Location/Qualifiers
      1..90

```

```

FEATURES
  source
    1..90

```

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0134E17"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/note="Vector: pWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      7 a  49 c  0 g  34 t
ORIGIN

Query Match      5.9%; Score 21; DB 12; Length 90;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 ggaggaggaggaggaggagg 349
|||||
Db 42 GGAGGGAGGGAGGGAGGAG 22

RESULT 38
AZ743191
LOCUS
DEFINITION RPI-24-146C22.TJ RPI-24 Mus musculus genomic clone RPI-24-146C22
, DNA sequence.
ACCESSION AZ743191
VERSION AZ743191.1 GI:12522805
KEYWORDS GSS.
SOURCE house mouse
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 120)
Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M., Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E., Russell,D., de Jong,P. and Fraser,C.M.
Mouse BAC End Sequences from Library RPI-24
Unpublished (1999)
Other GSSs: RPI-24-146C22.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPI-24. For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (http://www.choi.org/bacpac/orderingframe.htm). BAC end page: http://www.tigr.org/db/bac\_ends/mouse/bac\_end\_intro.html
Plate: 146 row: C column: 22
Seq primer: SP6
Class: BAC ends.

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0134E17"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/note="Vector: pWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      7 a  49 c  0 g  34 t
ORIGIN

Query Match      5.9%; Score 21; DB 12; Length 120;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 ggaggaggaggaggaggagg 349
|||||
Db 84 GGAGGGAGGGAGGGAGGAG 104

RESULT 39
AI053852/c
LOCUS
DEFINITION q170c08.x1 NCI-CGAP_Ow26 Homo sapiens CDNA clone IMAGE:1861838 3'
similar to contains element TAR1 TAR1 repetitive element ;, mRNA
sequence.
ACCESSION AI053852
VERSION AI053852.1 GI:3321639
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 145)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Monica Brown, M.D., Elise Kohn, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: David B. Krizman, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40m13 fwd. ET from Amersham.

Location/Qualifiers
1. .120
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPI-24-146C22"
/clone_lib="RPI-24"
/sex="Male"
/cell_type="Spleen/Brain"
/note="Vector: pTARBAC1; Site_1: BamHI; Site_2: BamHI; RPI-24 Mouse BAC Library produced by Pieter de Jong. The library was cloned in the pTARBAC1 cloning vector at the BamHI sites using MboI partially digested male C57BL/6J DNA."
BASE COUNT      42 a  16 c  46 g  16 t
ORIGIN

Query Match      5.9%; Score 21; DB 12; Length 120;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 ggaggaggaggaggaggagg 349
|||||
Db 84 GGAGGGAGGGAGGGAGGAG 104

RESULT 39
AI053852
LOCUS
DEFINITION q170c08.x1 NCI-CGAP_Ow26 Homo sapiens CDNA clone IMAGE:1861838 3'
similar to contains element TAR1 TAR1 repetitive element ;, mRNA
sequence.
ACCESSION AI053852
VERSION AI053852.1 GI:3321639
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 145)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Monica Brown, M.D., Elise Kohn, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: David B. Krizman, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40m13 fwd. ET from Amersham.

Location/Qualifiers
1. .145
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1861838"
/clone_lib="NCI-CGAP_Ow26"
/sex="female"
/tissue_type="papillary serous carcinoma"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: ovary; Vector: pAMPl; mRNA made from papillary serous ovarian carcinoma, cDNA made by oligo-dT priming. Directionally cloned. Size-selected on agarose gel, average insert size 500 bp. Primary library, non-amplified."
BASE COUNT      1 a  57 c  4 g  83 t
ORIGIN

```

Query Match 5.9%; Score 21; DB 9; Length 145;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 329 ggaggaggaggaggaggagg 349  
 |||||  
 Db 89 GGAGGGAGGGAGGAGGAAG 69

## RESULT 40

AZ091117  
 LOCUS RPCI-23-7H17.TJ RPCI-23 Mus musculus genomic clone RPCI-23-7H17,  
 DEFINITION DNA sequence.  
 ACCESSION AZ091117  
 VERSION AZ091117.1 GI:7733160  
 KEYWORDS GSS.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 161)  
 AUTHORS Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Akinret  
 ,B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.  
 and Fraser,C.M.  
 TITLE Mouse BAC End Sequences from Library RPCI-23  
 JOURNAL Unpublished (1999)  
 COMMENT Other\_GSSs: RPCI-23-7H17.TV  
 Contact: Shaying Zhao  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: szhao@tigr.org  
 Clones are derived from the mouse BAC library RPCI-23. For BAC  
 library availability, please contact Pieter de Jong  
 (pieter@dejong.med.buffalo.edu). Clones may be purchased from  
 BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)  
 or from Resea ch Genetics (info@resgen.com). BAC end page:  
 http://www.tigr.org/tdb/bac\_ends/mouse/bac\_end\_intro.html  
 Plate: 7 row: H column: 17  
 Seq primer: SP6  
 Class: BAC ends.

## FEATURES

Location/Qualifiers

1..161  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="RPCI-23-7H17"  
 /clone\_lib="RPCI-23"  
 /sex="Female"  
 /lab\_host="DH10B"

/note="Organ: Kidney/Brain; Vector: pBAC3.6; Site\_1:  
 EcoRI; Site\_2: EcoRI; Female C57BL/6J mouse kidney and/or  
 brain genomic DNA was isolated and partially digested  
 with a combination of EcoRI and EcoRI Methylase. Size  
 selected DNA was cloned into the pBAC3.6 vector at the  
 EcoRI sites. The ligation products were transformed into  
 DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 42 a 16 c 45 g 58 t  
 ORIGIN

Query Match 5.9%; Score 21; DB 12; Length 161;  
 Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 327 gagaggaggaggaggaggagg 347  
 |||||  
 Db 134 GAGAGGGAGGAGGAGGAAGA 154

## RESULT 41

AZ457379  
 LOCUS 191 bp DNA linear GSS 04-OCT-2000  
 DEFINITION 1M0260L10R Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
 clone UUGC1M0260L10 R, DNA sequence.  
 ACCESSION AZ457379  
 VERSION AZ457379.1 GI:10615504  
 KEYWORDS GSS.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 191)  
 AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly  
 ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.  
 and Wright,D., Weiss,R.  
 TITLE Mouse whole genome scaffolding with paired end reads from 10kb  
 plasmid inserts  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Robert B. Weiss  
 University of Utah Genome Center  
 University of Utah  
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
 84112, USA  
 Tel: 801 585 5606  
 Fax: 801 585 7177  
 Email: ddunn@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Plate: 0260 row: L column: 10  
 Seq primer: CACACAGGAAACAGCTATGACC  
 Class: plasmid ends  
 High quality sequence stop: 191.

## FEATURES

Location/Qualifiers

1..191

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="UUGC1M0260L10"

/clone\_lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab\_host="E. Coli strain XL10-Gold, Tl-resistant, F-"

/note="Vector: PWD42nv; Purified genomic DNA from M.  
 musculus C57BL/6J (male) was obtained from the Jackson  
 Laboratory Mouse DNA Resource  
 (http://www.jax.org/resources/documents/dnares/). The DNA  
 was hydrodynamically sheared by repeated passage through a  
 0.005 inch orifice at constant velocity. The sheared DNA  
 was blunt end-repaired with T4 DNA polymerase and T4  
 polynucleotide kinase. Adaptor oligonucleotides were  
 ligated to the blunt ends in high molar excess. The  
 adapted DNA was purified and size-selected for a 9.5 to  
 10.5 kb range using preparative agarose gel  
 electrophoresis. Vector DNA was prepared from a derivative  
 of pWD42 (gi|4732114|gb|AF129072.1), a copy-number  
 inducible derivative of plasmid R1. The vector was ligated  
 with adaptors complementary to the insert adaptors and  
 purified. The sheared, adapted mouse DNA was annealed to  
 adapted vector DNA, and transformed into  
 chemically-competent E. coli XL10-Gold (Stratagene) cells  
 and selected for ampicillin resistance."

BASE COUNT 46 a 55 c 44 g 46 t  
 ORIGIN

Query Match 5.9%; Score 21; DB 12; Length 191;  
 Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 327 gagaggaggaggaggaggagg 347  
 |||||



```

Db 169 GAGGAGGAGGAGGAGGAAGGA 189

RESULT 42
LOCUS A2656509 212 bp DNA linear GSS 14-DEC-2000
DEFINITION IM0532002F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0532002 F, DNA sequence.
ACCESSION A2656509
VERSION A2656509.1 GI:11793655
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 212)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.
and Wright,D., Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 309, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunne@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0532 row: 0 column: 02
Seq primer: CGTTGTAACGAGCGGCAGT
Class: plasmid ends
High quality sequence stop: 212.
FEATURES
Location/Qualifiers
1..212
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0532002"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adaptored DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PWD42 (gi14732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adaptored mouse DNA was annealed to
adaptored vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
BASE COUNT 39 a 83 c 43 g 47 t
ORIGIN

Query Match 5.9%; Score 21; DB 12; Length 212;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 ggaggaggaggaggaggaggag 349
|||||
Db 207 GGAGGGAGGAGGAGGAAGGAAG 227

RESULT 44
LOCUS AZ513751
DEFINITION IM0360802F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0360802 F, DNA sequence.
ACCESSION AZ513751
VERSION AZ513751.1 GI:10695067
KEYWORDS GSS.

```

SOURCE house mouse.  
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 294)

AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,  
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,  
M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.  
and Wright, D., Weiss, R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts

JOURNAL Unpublished (2000)

COMMENT Contact: Robert B. Weiss  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0360 row: B column: 02  
Seq primer: CGTTGTAACAGCGGCAGT  
Class: plasmid ends  
High quality sequence stop: 294.

FEATURES  
source  
1. .294  
Location/Qualifiers  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUGC1M0360B02"  
/clone\_lib="Mouse 10kb plasmid UUGC1M library"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, Ti-resistant, P-"  
/notes="Vector: PRD42nv; Purified genomic DNA from M.  
musculus C57BL/6J (male) was obtained from the Jackson  
Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA  
was hydrodynamically sheared by repeated passage through a  
0.005 inch orifice at constant velocity. The sheared DNA  
was blunt end-repaired with T4 DNA polymerase and T4  
polynucleotide kinase. Adaptor oligonucleotides were  
ligated to the blunt ends in high molar excess. The  
adapted DNA was purified and size-selected for a 9.5 to  
10.5 kb range using preparative agarose gel  
electrophoresis. Vector DNA was prepared from a derivative  
of pWB42 (gil4732114|gb|AF129072.1), a copy-number  
inducible derivative of plasmid R1. The vector was ligated  
with adaptors complementary to the insert adaptors and  
purified. The sheared, adapted mouse DNA was annealed to  
adapted vector DNA, and transformed into  
chemically-competent E. coli XL10-Gold (Stratagene) cells  
and selected for ampicillin resistance."

BASE COUNT 131 a 38 c 79 g 46 t  
ORIGIN

Query Match 5.9%; Score 21; DB 12; Length 294;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 ggaggaggaggaggaggaggagg 349  
|||||

Db 188 GGAGGGAGGGAGGGAGGAAG 208

RESULT 45  
B50482/c  
LOCUS B50482 319 bp DNA linear GSS 20-JUN-1998  
DEFINITION CIT-HSP-437M12.TP CIT-HSP Homo sapiens genomic clone 437M12, DNA  
sequence.  
ACCESSION B50482  
VERSION B50482.1 GI:2602719

KEYWORDS GSS.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 319)

AUTHORS Kim, U.-J., Adams, M.D. and Simon, M.I.

TITLE Determination of clone end sequences of human Bacterial Artificial  
Chromosomes

JOURNAL Unpublished (1997)

COMMENT Other\_GSSs: CIT-HSP-437M12.TV  
Contact: Ung-Jin Kim  
Caltech Genome Research Lab  
California Institute of Technology  
Division of Biology, MS 147-75, Pasadena, CA 91125, USA  
Tel: 626 796 7066  
Fax: 626 395 4901  
Email: ung@ash.tree.caltech.edu  
Clones are available from Research Genetics (info@resgen.com). BAC  
end search page:  
http://www.tigr.org/tldb/hungen/bac\_end\_search/bac\_end\_search.html  
Seq primer: SP6  
Class: BAC ends.

FEATURES  
source  
1. .319  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="GDB:5398850"  
/db\_xref="taxon:9606"  
/clone="437M12"  
/clone\_lib="CIT-HSP"  
/sex="Male"  
/cell\_type="Sperm"  
/note="Vector: pBeloBAC11; Site\_1: HindIII; Site\_2:  
HindIII"

BASE COUNT 47 a 119 c 21 g 132 t  
ORIGIN

Query Match 5.9%; Score 21; DB 12; Length 319;  
Best Local Similarity 100.0%; Pred. No. 12;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 329 ggaggaggaggaggaggaggagg 349  
|||||

Db 218 GGAGGGAGGGAGGGAGGAAG 198

Search completed: September 20, 2002, 04:07:24  
Job time: 13778 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2002, 06:30:00 ; Search time 5250.46 Seconds  
(without alignments)  
880.830 Million cell updates/sec

Title: US-09-846-456-4

Perfect score: 221

Sequence: 1 gtaattgcgagcgagtgta.....aacacaaagtggaaacag 221

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 1797656 seqs, 10453268293 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :		GenEmbl.*	
1:	gb_ba.*	1:	gb_ba.*
2:	gb_hgt.*	2:	gb_hgt.*
3:	gb_in.*	3:	gb_in.*
4:	gb_om.*	4:	gb_om.*
5:	gb_ov.*	5:	gb_ov.*
6:	gb_pat.*	6:	gb_pat.*
7:	gb_ph.*	7:	gb_ph.*
8:	gb_pl.*	8:	gb_pl.*
9:	gb_pr.*	9:	gb_pr.*
10:	gb_ro.*	10:	gb_ro.*
11:	gb_sts.*	11:	gb_sts.*
12:	gb_sv.*	12:	gb_sv.*
13:	gb_un.*	13:	gb_un.*
14:	gb_vl.*	14:	gb_vl.*
15:	em_ba.*	15:	em_ba.*
16:	em_fun.*	16:	em_fun.*
17:	em_hum.*	17:	em_hum.*
18:	em_in.*	18:	em_in.*
19:	em_mu.*	19:	em_mu.*
20:	em_on.*	20:	em_on.*
21:	em_or.*	21:	em_or.*
22:	em_ov.*	22:	em_ov.*
23:	em_pat.*	23:	em_pat.*
24:	em_ph.*	24:	em_ph.*
25:	em_pl.*	25:	em_pl.*
26:	em_ro.*	26:	em_ro.*
27:	em_sts.*	27:	em_sts.*
28:	em_un.*	28:	em_un.*
29:	em_vl.*	29:	em_vl.*
30:	em_hgt_hum.*	30:	em_hgt_hum.*
31:	em_hgt_inv.*	31:	em_hgt_inv.*
32:	em_hgt_other.*	32:	em_hgt_other.*
33:	em_hgtGO_inv.*	33:	em_hgtGO_inv.*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
-----					



QY 61 ccgggctcggcagcagggcggggagctccggcagcaccacagagccgggttctcagggc 120  
|||||  
Db 987 CCGGGCTCGGCGAGCGAGCGGGGAGCTCGCGCACCAACAGAGCGGTTCTCAGGGC 1046  
|||||  
QY 121 gcttgcctctgtttttcccggttctgttttcccttctccggaagcttgcac 180  
|||||  
Db 1047 GCTTTGCTCTGTGTTTCCCGGTTCTGTTTCTCCCTTCTCCGGAAGCTTGTCAC 1106  
|||||  
QY 181 ggggtagagaaagagcagcacaacaaagtggaaaaacag 221  
|||||  
Db 1107 GGGGTAGGAGAAAGAGACGCAACAAAGTGGAAAAACAG 1147  
|||||  
RESULT 4  
AF258623S1 1167 bp DNA linear PRI 23-JUN-2000  
LOCUS Homo sapiens ATP binding cassette transporter 1 (ABCA1) gene,  
DEFINITION promoter and exon 1.  
ACCESSION AF258623  
VERSION AF258623.2 GI:8677405  
KEYWORDS 1 of 4  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1167)  
AUTHORS Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C.,  
Aouizerat,B.E., Fielding,C.J. and Kane,J.P.  
TITLE Analysis of hABC1 gene 5' end; additional peptide sequence,  
JOURNAL promoter region, and four polymorphisms  
REFERENCE Biochem. Biophys. Res. Commun. 271 (2000) In press  
AUTHORS Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C.,  
Aouizerat,B.E., Fielding,C.J. and Kane,J.P.  
TITLE Direct Submission  
JOURNAL Submitted (20-APR-2000) Cardiovascular Research Institute,  
University of California, San Francisco, 505 Parnassus Avenue, San  
Francisco, CA 94143-0130, USA  
REFERENCE 3 (bases 1 to 1167)  
AUTHORS Pullinger,C.R., Hakamata,H., Duchateau,P.N., Eng,C.,  
Aouizerat,B.E., Fielding,C.J. and Kane,J.P.  
TITLE Direct Submission  
JOURNAL Submitted (23-JUN-2000) Cardiovascular Research Institute,  
University of California, San Francisco, 505 Parnassus Avenue, San  
Francisco, CA 94143-0130, USA  
REMARK Sequence update by submitter  
COMMENT On Jun 23, 2000 this sequence version replaced gi:7769713.  
FEATURES  
Source  
1..1167  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="9"  
/map="9q31"  
224..844  
/gene="ABCA1"  
845..1147  
/number=1  
promoter  
exon  
BASE COUNT 278 a 313 c 328 g 248 t  
ORIGIN  
Query Match 100.0%; Score 221; DB 9; Length 1167;  
Best Local Similarity 100.0%; Pred. No. 5.2e-118;  
Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 gtaattgcagcagagtagtgaggcgggaccccgagcagcagccgaccttctctc 60  
|||||  
Db 927 GTAAATGGAGCGAGAGTAGTGGGCGGGACCCGACGAGCGGACCCCTTCTCTC 986  
|||||  
QY 61 ccgggctcggcagcagggcggggagctccggcagcaccacagagccgggttctcagggc 120  
|||||

Db 987 CCGGGCTCGGCGAGCGAGCGGGGAGCTCGCGCACCAACAGAGCGGTTCTCAGGGC 1046  
QY 121 gcttgcctctgtttttcccggttctgttttcccttctccggaagcttgcac 180  
|||||  
Db 1047 GCTTTGCTCTGTGTTTCCCGGTTCTGTTTCTCCCTTCTCCGGAAGCTTGTCAC 1106  
|||||  
QY 181 ggggtagagaaagagcagcacaacaaagtggaaaaacag 221  
|||||  
Db 1107 GGGGTAGGAGAAAGAGACGCAACAAAGTGGAAAAACAG 1147  
|||||  
RESULT 5  
AX253452 3231 bp DNA linear PAT 06-FEB-2002  
LOCUS AX351029  
DEFINITION Sequence 1 from Patent WO0183746.  
ACCESSION AX351029  
VERSION AX351029.1 GI:18616385  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (sites)  
AUTHORS Rosier-Montus,M.F., Prades,C., Lemoline,C., Naudin,L., Deneffe,P.,  
Brewer,B., Duverger,N., Remaley,A. and Santamarina-Fojo,S.  
TITLE Regulatory nucleic acid sequences of the abcl gene  
JOURNAL Patent: WO 0183746-A 1 08-NOV-2001;  
Aventis Pharma S.A. (FR)  
FEATURES  
Location/Qualifiers  
source  
1..3231  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 809 a 773 c 876 g 773 t  
ORIGIN  
Query Match 100.0%; Score 221; DB 6; Length 3231;  
Best Local Similarity 100.0%; Pred. No. 5.3e-118;  
Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 gtaattgcagcagagtagtgaggcgggaccccgagcagcagccgaccttctctc 60  
|||||  
Db 2894 GTAATTCGAGCAGAGTAGTGGGCGGGACCCGACGAGCGGACCCCTTCTCTC 2953  
|||||  
QY 61 ccgggctcggcagcagggcggggagctccggcagcaccacagagccgggttctcagggc 120  
|||||  
Db 2954 CCGGGCTCGGCGAGCGGAGCGGGGAGCTCCGGCCACCAACAGAGCGGTTCTCAGGGC 3013  
|||||  
QY 121 gcttgcctctgtttttcccggttctgttttcccttctccggaagcttgcac 180  
|||||  
Db 3014 GCTTTGCTCTGTGTTTCCCGGTTCTGTTTCTCCCTTCTCCGGAAGCTTGTCAC 3073  
|||||  
QY 181 ggggtagagaaagagcagcacaacaaagtggaaaaacag 221  
|||||  
Db 3074 GGGGTAGGAGAAAGAGACGCAACAAAGTGGAAAAACAG 3114  
|||||  
RESULT 6  
AX253452 7260 bp DNA linear PAT 10-OCT-2001  
LOCUS AX253452  
DEFINITION Sequence 3 from Patent WO0170810.  
ACCESSION AX253452  
VERSION AX253452.1 GI:16073979  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 7260)  
AUTHORS Schmitz,G. and Bodzioch,M.  
TITLE Atp binding cassette transporter 1 (abcl) gene polymorphisms and  
uses thereof for the diagnosis and treatment of lipid,  
cardiovascular or inflammatory disorders











AF287262.1	GI:13876612				
VERSION	human.				
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
MEDLINE					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
FEATURES					
Source					
mRNA					
gene					
exon					
exon					
CDS					

```
/gene="ABCA1"
/number=15
/number=135898
/number=16
/number=137163
/number=17
/number=138391
/number=18
/number=140179
/number=19
/number=141340
/number=20
/number=142777
/number=21
/number=143123
/number=22
/number=143963
/number=23
/number=145441
/number=24
/number=146657
/number=25
/number=148527
/number=26
/number=148771
/number=27
/number=150280
/number=28
/number=152078
/number=29
/number=153438
/number=30
/number=156757
/number=31

Query Match
Best Local Similarity 100.0%; Score 221; DB 9; Length 201144;
Matches 221; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gtaattgcgagcagagtgagtgaggccggaccgcagagccgagccgacccctctctc 60
Db 33931 GTAATTCGAGCGAGAGTGAGTGGCGCGGACCCGACGAGCGGACCCCTTCTCTC 33990
QY 61 ccggggtcgcgagcagcgaggcgagtcgcgcgacacacagagccggttctcagggc 120
Db 33991 CCGGGTTCGCGAGCGGCGGCGGAGCTCCGCGCACCAACAGAGCCGGTTCTCAGGGC 34050
QY 121 gctttgctctgtttttcccggttctgttttctcccttctccggaaggttgtcaaa 180
Db 34051 GCITTTGCTCTGTTTTTCCCGGTTCTGTTTTCTCCCTTCTCCGGAAGGCTTGTCAA 34110
QY 181 ggggtagagaaagagacgcaaacacaaaagtggaaaacag 221
Db 34111 GGGGTAGAGAAAGAGACGCAACACAAAAGTGAAAAACAG 34151

RESULT 11
AK022254
LOCUS Homo sapiens cDNA FLJ12192 fis, clone MAMMA1000851. 1750 bp mRNA linear PRI 29-SEP-2000
DEFINITION AK022254
ACCESSION AK022254
VERSION AK022254.1 GI:10433612
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens Mammary gland cDNA to mRNA, clone_lib:MAMMAL1000851.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Takahashi-Fujii,A., Hara,H.,
Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R., Takeuchi,K.,
Arita,M., Nabekura,T., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J.,
Wakamatsu,A., Nakamura,Y., Nagahara,K., Masuho,Y. and Oshima,A.
NEDO human cDNA sequencing project
Unpublished (2000)
REFERENCE 2 (bases 1 to 1750)
AUTHORS Isogai,T. and Otsuki,T.
DIRECT SUBMISSION
JOURNAL Submitted (23-AUG-2000) to the DDBJ/EMBL/GenBank databases. Takao
ISOGAI, Helix Research Institute, Genomics Laboratory; 1532-3 Yana,
Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp,
Tel:81-438-52-3951, Fax:81-438-52-3952)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing and clone selection;
Helix Research Institute (supported by Japan Key Technology Center
etc.) and Department of Virology, Institute of Medical Science,
University of Tokyo.
FEATURES
Location/Qualifiers
1..1750
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MAMMA1000851"
/clone_lib="MAMMAL"
/tissue_type="Mammary gland"
/note="Cloning vector: pME18SFL3"
BASE COUNT 291 a 489 c 586 g 384 t
ORIGIN
```

```
QY 3 aattgcgagcagagtgagtgaggccggaccgcagagccgagccgacccctctctccc 62
Db 1 AATTGCGAGCGAGAGTGAGTGGCGCGGACCCGCGAGAGCGGACCCCTTCTCTCCC 60
QY 63 ggggtcgcgagcagcgaggcgagtcgcgcgacacacagagccggttctcaggcgcc 122
Db 61 GGGGTTCGCGAGCGGCGGCGGAGCTCCGCGCACCAACAGAGCGGTTCTCAGGGCGC 120
QY 123 ttgtctctgtttttcccggttctgttttctcccttctccggaaggttgtcaagg 182
Db 121 TTGTCTCTGTTGTTTTTCCCGGTTCTGTTTTCTCCCTTCTCCGGAAGGTTGTCAAG 180
QY 183 ggtagagaaagagacgcaaacacaaaagtggaaaacag 221
Db 181 GGTAGAGAAAGAGACGCAACACAAAAGTGAAAAACAG 219

RESULT 12
AK024328
LOCUS Homo sapiens cDNA FLJ14266 fis, clone PLACE1002437, highly similar
DEFINITION AK024328
ACCESSION AK024328
VERSION AK024328.1 GI:10436685
```



Query Match	92.8%;	Score 205;	DB 6;	Length 9854;
Best Local Similarity	100.0%;	Pred. No. 1.2e-108;		
Matches	205;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
17	gtgaagtggccgggaccgcagagccagagccagccctctctccgagcgtgcgcaggg 76			
Ddb	1 GTGAGTGGGGCCGGACCGCAGAGCGGACCGCTTCTCTCCGGGCTTGTGCTTCCTTT 120			
77	caggcggggagctccgcaccacacagagccggtttctcaggggcgtttgctcctgttt 136			
Ddb	61 CAGGCGGGGAGCTCCGCGCAGCCACACAGAGCGCGTCTCAGGGCGCTTGTGCTTCCTTT 120			
137	tttcccggttcgtttttccctctctcgaagagcgttgcagggtagagaaagag 196			
Ddb	121 TTTCCCGGTTCTGTTTTCTCCCTTCTCGGAAGCTTGTCAAGGGGTAGGAAAGAG 180			
197	acgcaaacacaaagtggaaaacag 221			
Ddb	181 ACGCAACACAAAGTGGAAACAG 205			

RESULT 15	AC021246	69570 bp	DNA	linear	HTG 13-JUL-2000
LOCUS	AC021246	Homo sapiens clone RP11-1N10, LOW-PASS SEQUENCE SAMPLING.			
DEFINITION	AC021246				
ACCESSION	AC021246.2	GI:9119882			
VERSION	HTG; HTGS_PHASE0.				
KEYWORDS	human.				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 69570)				
AUTHORS	Birren,B., Linton,L., Nusbaum,C. and Lander,E.				
TITLE	Homo sapiens chromosome, clone RP11-1N10				
REFERENCE	2 (bases 1 to 69570)				
AUTHORS	Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F., Boguslavsky,L., Bouknight,B., Brown,A., Burkett,G., Castle,A., Choepel,X., Collangelo,M., Collins,S., Collymore,A., Cooke,P., DeArrellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J., Ferreira,P., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Johnson,K., Jones,C., Kann,L., Karatas,A., Klein,J., Landers,T., Lechovsky,J., Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D., Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.				
TITLE	Direct Submission				
JOURNAL	Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA				
COMMENT	On Jul 13, 2000 this sequence version replaced gi:6705871. All repeats were identified using RepeatMasker: Smit,A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html				
	----- Genome Center				
	Center: Whitehead Institute/ MIT Center for Genome Research				
	Center code: WIBR				
	Web site: http://www-seq.wi.mit.edu				
	Contact: sequence_submissions@genome.wi.mit.edu				
	----- Project Information				
	Center project name: L2512				
	Center clone name: L_N10				
	-----				
	* NOTE: This record contains 73 individual				
	* sequencing reads that have not been assembled into				

30556	31410:	contig of 855 bp	in length
31411	31510:	gap of 100 bp	
31511	32368:	contig of 858 bp	in length
32369	32468:	gap of 100 bp	
32469	33312:	contig of 844 bp	in length
33313	33412:	gap of 100 bp	
33413	34268:	contig of 856 bp	in length
34269	34368:	gap of 100 bp	
34369	35204:	contig of 836 bp	in length
35205	35304:	gap of 100 bp	
35305	36156:	contig of 852 bp	in length
36157	36256:	gap of 100 bp	
36257	37118:	contig of 872 bp	in length
37129	37228:	gap of 100 bp	
37229	38083:	contig of 855 bp	in length
38084	38183:	gap of 100 bp	
38184	39031:	contig of 848 bp	in length
39032	39131:	gap of 100 bp	
39132	40006:	contig of 875 bp	in length
40007	40106:	gap of 100 bp	
40107	40967:	contig of 861 bp	in length
40968	41067:	gap of 100 bp	
41068	41913:	contig of 846 bp	in length
41914	42013:	gap of 100 bp	
42014	42824:	contig of 811 bp	in length
42825	42924:	gap of 100 bp	
42925	43776:	contig of 852 bp	in length
43777	43876:	gap of 100 bp	
43877	44752:	contig of 876 bp	in length
44753	44852:	gap of 100 bp	
44853	45724:	contig of 872 bp	in length
45725	45824:	gap of 100 bp	
45825	46643:	contig of 819 bp	in length
46644	46743:	gap of 100 bp	
46744	47599:	contig of 856 bp	in length
47600	47699:	gap of 100 bp	
47700	48551:	contig of 852 bp	in length
48552	48651:	gap of 100 bp	
48652	49485:	contig of 834 bp	in length
49486	49586:	gap of 100 bp	
49586	50440:	contig of 855 bp	in length
50441	50540:	gap of 100 bp	
50541	51404:	contig of 864 bp	in length
51405	51504:	gap of 100 bp	
51505	52372:	contig of 868 bp	in length
52373	52472:	gap of 100 bp	
52473	53328:	contig of 856 bp	in length
53329	53428:	gap of 100 bp	
53429	54268:	contig of 840 bp	in length
54269	54368:	gap of 100 bp	
54369	55229:	contig of 861 bp	in length
55230	55329:	gap of 100 bp	
55330	56197:	contig of 868 bp	in length
56198	56297:	gap of 100 bp	
56298	57163:	contig of 866 bp	in length
57164	57263:	gap of 100 bp	
57264	58130:	contig of 867 bp	in length
58131	58230:	gap of 100 bp	
58231	59082:	contig of 852 bp	in length
59083	59182:	gap of 100 bp	
59183	60020:	contig of 838 bp	in length
60021	60120:	gap of 100 bp	
60121	60983:	contig of 863 bp	in length
60984	61083:	gap of 100 bp	
61084	61935:	contig of 852 bp	in length
61936	62035:	gap of 100 bp	
62036	62866:	contig of 831 bp	in length
62867	62966:	gap of 100 bp	
62967	63827:	contig of 861 bp	in length
63828	63927:	gap of 100 bp	
63928	64783:	contig of 856 bp	in length
64784	64883:	gap of 100 bp	
64884	65740:	contig of 857 bp	in length

*	65741	65840:	gap of	100 bp
*	65841	66684:	contig of 844 bp in length	
*	66685	66784:	gap of	100 bp
*	66785	67651:	contig of 867 bp in length	
*	67652	67751:	gap of	100 bp
 Query Match Best Local Similarity   91.0%; Score 201; DB 2; Length 69570; Matches 201; Conservative   0; Mismatches   0; Indels   0; Gaps   0;				
Qy	1	gtaattcgagcgcagagtga	tggggccggacccgcagaccgagccgagccgcccttcttc	60
Dd	41565	GTAATTGGCAGCGACAGTG	TGCGGGCCGGACCCGCAGAGCCGACGCCCTTCTCTC	41624
Qy	61	ccgggtcgcgccaggcgagg	cggggagctccgcgcaccaacagagccggtttctcagggc	120
Dd	41625	CCGGGTCTCGCAGCGCAGG	CAGCGCGGGAGCTCGCGCACCAACAGAGCCGTTTC	41684
Qy	121	gcttttgtacctgttttttc	cccgggttctgttttcccccttctccgaaaggcttgtaa	180
Dd	41685	GCTTTGTCCTCTGTGTTTT	TCCC CGGTTCTGTTCCTTCCTTCCTCGAAGGCTTGTCAA	41744
Qy	181	ggggttaggagaaaagacgc	a 201	
Dd	41745	GGGGTAGGAAGAAGACGCA	41765	
 RESULT 16 AX092589 LOCUS AX092589 183999 bp DNA linear PAT 21-MAR-2001 DEFINITION Sequence 1 from Patent WO0115676. ACCESSION AX092589 VERSION AX092589.1 GI:13444647 KEYWORDS . SOURCE human. ORGANISM Homo sapiens  REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. AUTHORS 1 (bases 1 to 183999) TITLE Hayden,M.R., Brooks-Wilson,A.R., Pimstone,S.N. and Clee,S.M. Compositions and methods for modulating hdl cholesterol and triglyceride levels JOURNAL Patent: WO 0115676-A 1 08-MAR-2001; University of British Columbia (CA) ; Xenon Genetics Inc. (CA) FEATURES Location/Qualifiers source 1..183999 /organism="Homo sapiens" /db_xref="taxon:9606"				
BASE COUNT	49549	a	37944 c	41170 g 54950 t 386 others
ORIGIN				
 Query Match Best Local Similarity   91.0%; Score 201; DB 6; Length 183999; Matches 201; Conservative   0; Mismatches   0; Indels   0; Gaps   0;				
Qy	1	gtaattcgagcgcagagtga	tggggccggacccgcagaccgagccgagccgcccttcttc	60
Dd	28790	GTAATTGCGACGCGAGAGTG	AGTGGCGCGGACCCGCAGAGCCGACGCCCTTCTCTC	28849
Qy	61	ccgggtcgcgccaggcgagg	cggggagctccgcgcaccaacagagccggtttctcagggc	120
Dd	28850	CCGGGTCTCGCAGCGGCGAG	GCGGGGAGCTCGCGCACCAACAGAGCCGTTTC	28909
Qy	121	gcttttgtacctgttttttc	cccgggttctgttttcccccttctccgaaaggcttgtaa	180
Dd	28910	GCTTTGTCCTCTGTGTTTT	TCCC CGGTTCTGTTCCTTCCTTCCTCGAAGGCTTGTCAA	28969
Qy	181	ggggttaggagaaaagacgc	a 201	
Dd	28970	GGGGTAGGAAGAAGACGCA	28990	



```

RESULT 17
AX060713
LOCUS      10442 bp      DNA      linear      PAT 22-JAN-2001
DEFINITION Sequence 1 from Patent WO0078972.
ACCESSION AX060713
VERSION    AX060713.1 GI:12406103
KEYWORDS   human.
SOURCE     Homo sapiens
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 10442)
AUTHORS    Lawn,R.M., Wade,D. and Garvin,M.
TITLE      Regulation with binding cassette transporter protein abcl
JOURNAL    Patent: WO 0078972-A 1 28-DEC-2000;
           CV THERAPEUTICS, INC. (US)
FEATURES   Location/Qualifiers
            source
              1..10442
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
BASE COUNT 2898 a 2297 c 2408 g 2835 t      4 others
ORIGIN

Query Match      89.1%; Score 197; DB 6; Length 10442;
Best Local Similarity 100.0%; Pred. No. 5.9e-104;
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 ggcgggacccgcagagccgagccgacctctctccgggctgcgagcgagggcg 84
|||||
Db 1  GCCGGGACCCGACAGCCGAGCCGACCCCTCTCTCCGGGCTCGGAGCGGCGGCG 60
|||||

QY 85 ggagctccgcgcacacagagccggttctcagggcgttgcctcttctttccccc 144
|||||
Db 61 GGAGCTCCGGCGACACAGAGCGGCTTCTCAGGGCGCTTGTCTCTCTTCTTCCCG 120
|||||

QY 145 gttctgtttctccctctccggaagccttgcacaggttagagagaaagacgcaaac 204
|||||
Db 121 GTTCTGTCTTCTCCCTCTCTCCGAGGCTTGTCAAGGGTAGGAGAAAGACGCAAC 180
|||||

QY 205 acaaaagtggaaacag 221
|||||
Db 181 ACAAAAGTGGAACAG 197
|||||

RESULT 19
AF285167      10442 bp      mRNA      linear      PRI 09-AUG-2000
LOCUS      Homo sapiens ATP-binding cassette transporter 1 (ABCA1) mRNA,
DEFINITION complete cds.
ACCESSION    AF285167
VERSION      AF285167.1 GI:9755158
KEYWORDS     human.
SOURCE       Homo sapiens
ORGANISM     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1 (bases 1 to 10442)
AUTHORS      Schwartz,K., Lawn,R.M. and Wade,D.P.
TITLE        ABCA1 gene expression and apoA-I-mediated cholesterol efflux are
           regulated by LXR
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 10442)
AUTHORS      Lawn,R.M., Wade,D.P., Garvin,M.R., Wang,X., Schwartz,K.,
           Porter,J.G., Seilhamer,J.J., Vaughan,A.M. and Oram,J.F.
TITLE        Direct Submision
JOURNAL      Submitted (06-JUL-2000) Discover Research, CV Therapeutics Inc.,
           3172 Porter Drive, Palo Alto, CA 94304, USA
FEATURES     Location/Qualifiers
              1..10442
              /organism="Homo sapiens"
              /db_xref="taxon:9606"
              /chromosome="9"
              /map="9q31"
              /cell_type="fibroblast"
              /tissue_type="skin"
              1..10442
              /gene="ABCA1"
              291..7076
              /gene="ABCA1"
              /codon_start=1
              /product="ATP-binding cassette transporter 1"
              /protein_id="AAF98175.1"
              /db_xref="GI:9755159"
              /translation="MACWPQLRLLLKWLKLTERRQTCQLLELVAMPFLIFLILSVRL
              SYPEOHECHEPNKPAGTLPWQGLICNANPCFXYPTCPAGVGVGNFNKSV
              ARLFSDARLLLYSKQTSMDMKVLTQQIKSSNLKLOFLVDNETFSCFLVH
              NLSLPKSTVDKMLRADVILHKVFLQGYQLHLTSLCNGSKSEEMQLQDQGVSELCGLP
              KEKLAARVLRNNMDILPLRTLNSTSPFSPKSLAEATKTLHLSTLAQELFSMR
              SWSMDROEVMTNVNSSSTQIYQAVSRVCGHPGGGLKIKSLNMYEDNNYKALF
              GNGTEDAETFDYNTTPYCNLDLNMKLESSPLRIWKALKPLVCKILVTPDTPT
              RQVAEYNTQELAVFHDLSGMHEELSPKIWTENSOEMDLVRLMDSRDNHFW
              COLDLMDTAQDIAVFLAKHPEDVQSSNGSVITWEAFNFTNQAIRISRMECVNIL
              KLEPATEWMLINKSMELLDERKFMAGIVFTGTPGSELPHHVKYIKRIMDINVERT
              NKIKDGYWDPGRADPFEDMYVWGQFAYLDQVVEQAIIRVLTGTEKTKTGYMQMPY
              PCYVDDIFLRVMSRMPMTLAWISVAVIKGIVYEKARLKTMRIMGLDNLILW
              FSWPTSSILPLLVSAGLLVTLKGLNLLPYSDPSVVFVFLSVFAVVTIQCFLISTLF
              SRANLAACGGIIIFTLYLPLVLCVWQDYGFTLKIFASLLSPVAFGFGCEYFALFE
              EQGIGVQNDNLFPESVEDGNLTSTISMMLFDFTLYGWTWTVYEAFFPGQYIPRW
              YFPCTKSYFGEESDEKSHPGSNQKRMSEICMEEEPETHLKGVSIGNLVKVRDGMKV
              AVDGLALNFYEGQITSLFLGHNGAGKTTTMSILTGLFPPTSGTAVILGKDIRSEMSTR

```

BASE COUNT ORIGIN	2898 a	2297 c	2408 g	2835 t	4 others
	ONLQVCPQHNVLDMLTVEBHIWFYARLKGKLSKHVKAEMQMALDVGLPSSKLKSKT				
	SHQSDGMRKLSVALAPVGGSKVYILDEPTGAPDPVYSRRGIIWELLIKYRGOSTRIILT				
	HHMEDADVLDRIATISHGKLCGVSSLFKNQGLGTYYLTIVKKDVSSSCRNSS				
	STVYILKEDSVQSSSDAGLSDHSDTUTIDVSAISNGLIRKHVSEARLIVEDIGHEL				
	TYVLPYEAKEAGAVELFIEDRLSDLGISSYGISETTILEEIFLKVAESGVDATIS				
	DGTLIPARRNRARFGDKQSCURPTEDDAADPNDSIDPESRETDILSGMDKGSGYQVK				
	GWKLTQQOQFVALLKRLDITARRSKGFFAQIVLPAFVFCIALVFSLIVPPFGYPSLE				
	LQPMWYNEQYTFWSSNDAPEGTGLELINALTKDPGFCTRCMEGNIPDTPCQAGEEW				
	TTAPVPTIMDLFQNGWNTMONPSPACQCSSDKIKMLPVCPPGACGLPDPORQONTA				
	DILQDLTGRNISDYLKTVYQIILAKLSKNKIWNERYGGFSLGVSNTQALPPSQEVN				
	DAIKQMKHLKLAKDSADRAFLNSLGRFMTGLDTRNNKVMFNKNGKHAISLFLNVIN				
	NAILLRANLQGENPSHYGTAFNHLNLTQKQSEVALMTSDVDSLVSICVIFAMSFV				
	PASGVPYLIQERYSKAKHLQFISGVKPVYIWNLSFWMDMNVVPAVLTVIIFCQQ				
	KSVSNSTNPVALALLLLYGWSITPLMWIPASVFEKIPSTAYVVLTVSNLFIGNGVA				
	TVYLELTQNLNNIINDLKLSVELIFPHCLGRLGIDMWKNQAMADALERFGNRPVS				
	PLSWDLVGRNLFMAVEGVVFFLTVLIQTRFIRPRPVNAKLUSDDEEDYVRERQ				
	RILUDGGQNDILEIKETIYLRKKRPADVRICVGPPEGCGLLGNAGKSGKSTFKM				
	LDGTTVITRGDPAFLNKSILSINTIYHQNMVYCPQDFATELLTGRHEVFEFALLRGV				
	PEKEVGGEWALRKGLGVYGBKYAGNYSGGNKRLKSTAMALIGGPPVYFLDEPTTG				
	MDPKARRELFWNCAKSVKGRSVVLVTSHSMECEALCTRMAIWNVNGPFCPLSGVQHLK				
	NRFGDGYTIIVRTAGSNPDLKPVQDFEGFLAPGVSILKEHRNMLOQLPSSGYSLSLARI				
	FTSLQSKSKRHLIEDYSVOTTLDUDQVFNFAKQDQSDDDHLKOLSLKHKNQTVVDVAVLT				
	FSFLQDSKVKRESY*				

STUDY AREA				
BASE COUNT	2898 a	2297 c	2408 g	2835 t
ORIGIN				
4 others				

Query Match 89.1%; Score 197; DB 9; Length 10442;

[illegible]

RESULT	20
AC021345/C	
LOCUS	AC021345
DEFINITION	Homo sapiens clone RP11-24J9, LOW-PASS SEQUENCE SAMPLING.
ACCESSION	AC021345
VERSION	AC021345.2 GI:9130845
KEYWORDS	HTG; HTGS_PHASEO.
SOURCE	human.
ORGANISM	Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 90698)  
Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
Homo sapiens, clone RP11-24J9  
Unpublished  
2 (bases 1 to 90698)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F., Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,I., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Landers,T., Lehoczeky,J., Levine,R., Lieu,C., Liu,G., Locke,K.,



```

BASE COUNT      2906 a  2305 c  2416 g  2843 t      4 others
ORIGIN

Query Match      85.1%; Score 188; DB 6; Length 10474;
Best Local Similarity 100.0%; Pred. No. 1.1e-98;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  34  ccgcagagccgagccgacctctctcccggtctgcggcgagcgagcgagcgagctccg 93
      |||
Db  42  CCGCAGAGCCGAGCCGACCTTCTCTCCGGCTGCGCAGGAGCGAGCGGGAGCTCCG 101
      |||

QY  94  cgcaccaacagagcggtctcagggcgcttgctcctgtttttcccggttctgttt 153
      |||
Db  102  CGCACCAACAGAGCGGTTCTCAGGGGCTTGCTCTCTGTTTTTCCCGGTTCTGTTT 161

QY  154  tctccctctccgaaagctgtcaagggtagaagaaagagacgcaacacaaaagt 213
      |||
Db  162  TCTCCCTTCTCCGGAAGGCTTGTCAAGGGGTAGGAAAGAGACGCAACACAAAAGTG 221

QY  214  gaaaacag 221
      |||
Db  222  GAAACAG 229

RESULT 22
AX060721
LOCUS      AX060721
DEFINITION Sequence 9 from Patent WO0078972.
ACCESSION  AX060721
VERSION     AX060721.1 GI:12406109
KEYWORDS   human.
SOURCE     Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 10474)
AUTHORS   Lawn,R.M., Wade,D., and Garvin,M.
TITLE     Regulation with binding cassette transporter protein abcl
JOURNAL   CV THERAPEUTICS, INC. (US)
FEATURES   Location/Qualifiers
           source
             1..10474
             /organism="Homo sapiens"
             /db_xref="taxon:9606"

BASE COUNT      2907 a  2304 c  2415 g  2844 t      4 others
ORIGIN

Query Match      85.1%; Score 188; DB 6; Length 10474;
Best Local Similarity 100.0%; Pred. No. 1.1e-98;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  34  ccgcagagccgagccgacctctctcccggtctgcggcgagcgagcgagcgagctccg 93
      |||
Db  42  CCGCAGAGCCGAGCCGACCTTCTCTCCGGCTGCGCAGGAGCGAGCGGGAGCTCCG 101
      |||

QY  94  cgcaccaacagagcggtctcagggcgcttgctcctgtttttcccggttctgttt 153
      |||
Db  102  CGCACCAACAGAGCGGTTCTCAGGGGCTTGCTCTCTGTTTTTCCCGGTTCTGTTT 161

QY  154  tctccctctccgaaagctgtcaagggtagaagaaagagacgcaacacaaaagt 213
      |||
Db  162  TCTCCCTTCTCCGGAAGGCTTGTCAAGGGGTAGGAAAGAGACGCAACACAAAAGTG 221

QY  214  gaaaacag 221
      |||
Db  222  GAAACAG 229

RESULT 23
AX060898
LOCUS      AX060898
DEFINITION Sequence 9 from Patent WO0078971.
ACCESSION  AX060898
VERSION     AX060898.1 GI:12406275
KEYWORDS   human.
SOURCE     Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 10474)
AUTHORS   Lawn,R.M., Wade,D., Oram,J.F. and Garvin,M.
TITLE     Atp binding cassette transporter protein abcl polypeptides
JOURNAL   CV THERAPEUTICS, INC. (US)
FEATURES   Location/Qualifiers
           source
             1..10474
             /organism="Homo sapiens"
             /db_xref="taxon:9606"

BASE COUNT      2907 a  2304 c  2415 g  2844 t      4 others
ORIGIN

Query Match      85.1%; Score 188; DB 6; Length 10474;
Best Local Similarity 100.0%; Pred. No. 1.1e-98;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  34  ccgcagagccgagccgacctctctcccggtctgcggcgagcgagcgagcgagctccg 93
      |||
Db  42  CCGCAGAGCCGAGCCGACCTTCTCTCCGGCTGCGCAGGAGCGAGCGGGAGCTCCG 101
      |||

```

```

DEFINITION      Sequence 7 from Patent WO0078971.
ACCESSION       AX060898
VERSION         AX060898.1 GI:12406275
KEYWORDS        human.
SOURCE          Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE       1 (bases 1 to 10474)
AUTHORS        Lawn,R.M., Wade,D., Oram,J.F. and Garvin,M.
TITLE          Atp binding cassette transporter protein abcl polypeptides
JOURNAL        Patent: WO 0078971-A 7 28-DEC-2000;
               CV THERAPEUTICS, INC. (US)
FEATURES        Location/Qualifiers
               source
                 1..10474
                 /organism="Homo sapiens"
                 /db_xref="taxon:9606"

BASE COUNT      2906 a  2305 c  2416 g  2843 t      4 others
ORIGIN

Query Match      85.1%; Score 188; DB 6; Length 10474;
Best Local Similarity 100.0%; Pred. No. 1.1e-98;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  34  ccgcagagccgagccgacctctctcccggtctgcggcgagcgagcgagcgagctccg 93
      |||
Db  42  CCGCAGAGCCGAGCCGACCTTCTCTCCGGCTGCGCAGGAGCGAGCGGGAGCTCCG 101
      |||

QY  94  cgcaccaacagagcggtctcagggcgcttgctcctgtttttcccggttctgttt 153
      |||
Db  102  CGCACCAACAGAGCGGTTCTCAGGGGCTTGCTCTCTGTTTTTCCCGGTTCTGTTT 161

QY  154  tctccctctccgaaagctgtcaagggtagaagaaagagacgcaacacaaaagt 213
      |||
Db  162  TCTCCCTTCTCTCCGGAAGGCTTGTCAAGGGGTAGGAAAGAGACGCAACACAAAAGTG 221

QY  214  gaaaacag 221
      |||
Db  222  GAAACAG 229

RESULT 24
AX060900
LOCUS      AX060900
DEFINITION Sequence 9 from Patent WO0078971.
ACCESSION  AX060900
VERSION     AX060900.1 GI:12406276
KEYWORDS   human.
SOURCE     Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 10474)
AUTHORS   Lawn,R.M., Wade,D., Oram,J.F. and Garvin,M.
TITLE     Atp binding cassette transporter protein abcl polypeptides
JOURNAL   Patent: WO 0078971-A 9 28-DEC-2000;
               CV THERAPEUTICS, INC. (US)
FEATURES   Location/Qualifiers
           source
             1..10474
             /organism="Homo sapiens"
             /db_xref="taxon:9606"

BASE COUNT      2907 a  2304 c  2415 g  2844 t      4 others
ORIGIN

Query Match      85.1%; Score 188; DB 6; Length 10474;
Best Local Similarity 100.0%; Pred. No. 1.1e-98;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  34  ccgcagagccgagccgacctctctcccggtctgcggcgagcgagcgagcgagctccg 93
      |||
Db  42  CCGCAGAGCCGAGCCGACCTTCTCTCCGGCTGCGCAGGAGCGAGCGGGAGCTCCG 101
      |||

```

```
QY 94 cgcacacacagcgcgttctcagggcgttctgtctctgtttttttcccccgttctgttt 153
|||||
Db 102 CGCACAAACAGAGCGGTCTCAGGCGCTTGTGCTCTGTTGTTTTCGCCGGTCTCTGTTT 161
|||||
QY 154 tctccctctcgcgaagcgttgcagggtagagaaagagcgcacacacacaaagt 213
|||||
Db 162 TCTCCCTCTCTCCGAAGCTTGTCAAGGGGTAGGAGAAAGAGACGCAACACAAAAGTG 221
|||||
QY 214 gaaaacag 221
|||||
Db 222 GAAACAG 229

RESULT 25
AX127764
LOCUS AX127764 446 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 3 from Patent WO0130848.
ACCESSION AX127764
VERSION AX127764.1 GI:14134411
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 446)
AUTHORS Deneffe,P., Rosier-Montus,M.F., Arnould-Reguigne,I., Prades,C.,
Naudin,L., Lemoine,C., Duverger,N., Jaye,M., Searfoss,G.H.,
Remaley,A., Brewer,H.B. and Dean,M.
TITLE Nucleic acids of the human abcl gene and their therapeutic and
diagnostic application
JOURNAL Patent: WO 0130848-A 3 03-MAY-2001;
Aventis Pharma S.A. (FR)
FEATURES
source Location/Qualifiers
1..446
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Oligonucleotide Primer"
BASE COUNT 96 a 123 c 112 g 115 t
ORIGIN
Query Match 41.6%; Score 92; DB 6; Length 446;
Best Local Similarity 100.0%; Pred. No. 1.6e-42;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 130 cttgtttttcccccgttctgttttcccttccgcgaagcgttgcaggggtagga 189
|||||
Db 1 CTTGTTTTTCCCGGTTCTGTTTCTCCCTTCTCCGGAAGGCTTGTCAAGGGGTAGGA 60
|||||
QY 190 gaaagagcgcacacacacaaagtggaaaacag 221
|||||
Db 61 GAAAGAGACGCAACACAAAGTGGAACACAG 92
|||||
RESULT 26
AX139751
LOCUS AX139751 446 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 3 from Patent EP1096012.
ACCESSION AX139751
VERSION AX139751.1 GI:14275333
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 446)
AUTHORS Deneffe,P., Rosier-Montus,M.F., Arnould-Reguigne,I., Prades,C.,
Naudin,L., Lemoine,C., Duverger,N., Jaye,M., Searfoss,G.H.,
Remaley,A., Brewer,H.B. and Dean,M.
TITLE Nucleic acids of the human abcl gene and their therapeutic and
diagnostic application
JOURNAL Patent: EP 1096012-A 3 02-MAY-2001;
Aventis Pharma S.A. (FR)
FEATURES
source Location/Qualifiers
1..446
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Oligonucleotide Primer"
BASE COUNT 96 a 123 c 112 g 115 t
ORIGIN
Query Match 41.6%; Score 92; DB 6; Length 446;
Best Local Similarity 100.0%; Pred. No. 1.6e-42;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 130 cttgtttttcccccgttctgttttcccttccgcgaagcgttgcaggggtagga 189
|||||
Db 1 CTTGTTTTTCCCGGTTCTGTTTCTCCCTTCTCCGGAAGGCTTGTCAAGGGGTAGGA 60
|||||
QY 190 gaaagagcgcacacacacaaagtggaaaacag 221
|||||
Db 61 GAAAGAGACGCAACACAAAGTGGAACACAG 92
|||||
```

```
source 1..446
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Oligonucleotide Primer"
BASE COUNT 96 a 123 c 112 g 115 t
ORIGIN
Query Match 41.6%; Score 92; DB 6; Length 446;
Best Local Similarity 100.0%; Pred. No. 1.6e-42;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 130 cttgtttttcccccgttctgttttcccttccgcgaagcgttgcaggggtagga 189
|||||
Db 1 CTTGTTTTTCCCGGTTCTGTTTCTCCCTTCTCCGGAAGGCTTGTCAAGGGGTAGGA 60
|||||
QY 190 gaaagagcgcacacacacaaagtggaaaacag 221
|||||
Db 61 GAAAGAGACGCAACACAAAGTGGAACACAG 92
|||||
RESULT 27
AX127830
LOCUS AX127830 9741 bp DNA linear PAT 15-MAY-2001
DEFINITION Sequence 69 from Patent WO0130848.
ACCESSION AX127830
VERSION AX127830.1 GI:14134477
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 9741)
AUTHORS Deneffe,P., Rosier-Montus,M.F., Arnould-Reguigne,I., Prades,C.,
Naudin,L., Lemoine,C., Duverger,N., Jaye,M., Searfoss,G.H.,
Remaley,A., Brewer,H.B. and Dean,M.
TITLE Nucleic acids of the human abcl gene and their therapeutic and
diagnostic application
JOURNAL Patent: WO 0130848-A 69 03-MAY-2001;
Aventis Pharma S.A. (FR)
FEATURES
source Location/Qualifiers
1..9741
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2650 a 2180 c 2290 g 2620 t 1 others
ORIGIN
Query Match 41.6%; Score 92; DB 6; Length 9741;
Best Local Similarity 100.0%; Pred. No. 1.7e-42;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 130 cttgtttttcccccgttctgttttcccttccgcgaagcgttgcaggggtagga 189
|||||
Db 1 CTTGTTTTTCCCGGTTCTGTTTCTCCCTTCTCCGGAAGGCTTGTCAAGGGGTAGGA 60
|||||
QY 190 gaaagagcgcacacacacaaagtggaaaacag 221
|||||
Db 61 GAAAGAGACGCAACACAAAGTGGAACACAG 92
|||||
RESULT 28
AX139817
LOCUS AX139817 9741 bp DNA linear PAT 30-MAY-2001
DEFINITION Sequence 69 from Patent EP1096012.
ACCESSION AX139817
VERSION AX139817.1 GI:14275399
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 9741)
AUTHORS Deneffe,P., Rosier-Montus,M.F., Arnould-Reguigne,I., Prades,C.,
Naudin,L., Lemoine,C., Duverger,N., Jaye,M., Searfoss,G.H.,
Remaley,A., Brewer,H.B. and Dean,M.
TITLE Nucleic acids of the human abcl gene and their therapeutic and
diagnostic application
JOURNAL Patent: EP 1096012-A 3 02-MAY-2001;
Aventis Pharma S.A. (FR)
FEATURES
source Location/Qualifiers
1..9741
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 2650 a 2180 c 2290 g 2620 t 1 others
ORIGIN
Query Match 41.6%; Score 92; DB 6; Length 9741;
Best Local Similarity 100.0%; Pred. No. 1.7e-42;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 130 cttgtttttcccccgttctgttttcccttccgcgaagcgttgcaggggtagga 189
|||||
Db 1 CTTGTTTTTCCCGGTTCTGTTTCTCCCTTCTCCGGAAGGCTTGTCAAGGGGTAGGA 60
|||||
QY 190 gaaagagcgcacacacacaaagtggaaaacag 221
|||||
Db 61 GAAAGAGACGCAACACAAAGTGGAACACAG 92
|||||
```

AUTHORS Denefle, P., Rosier-Montus, M.F., Arnould-Reguigne, I., Prades, C., Naudin, L., Lemoine, C., Duverger, N., Jaye, M., searfoos Iii, G.H., Remaley, A., Brewer, H.B. and Dean, M.

TITLE Nucleic acids of the human abcl gene and their therapeutic and diagnostic application

JOURNAL Patent: EP 1056012-A 69 02-MAY-2001; Aventis Pharma S.A. (FR)

FEATURES  
source Location/Qualifiers  
1..9741  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 2650 a 2180 c 2290 g 2620 t 1 others  
ORIGIN

Query Match 41.6%; Score 92; DB 6; Length 9741;  
Best Local Similarity 100.0%; Pred. No. 1.7e-42;  
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 cttgtttttcccggttctgttttcccttcccttccggaagcttgcgaagggttagga 189  
|||||  
Db 1 CTTGTTTTTCCCGGTTCTGTTTCTCCCTTCTCCGAGAGGCTTGTCAGGGGTAGGA 60  
|||||

QY 190 gaaagagacgcaaacacaaaagtgaacaag 221  
|||||  
Db 61 GAAAGAGACGCAACACAAAAGTGGAACAG 92  
|||||

RESULT 29  
AX351038 9741 bp DNA linear PAT 06-FEB-2002  
LOCUS  
DEFINITION Sequence 10 from Patent WO0183746.  
ACCESSION AX351038  
VERSION AX351038.1 GI:18616393  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (sites)  
AUTHORS Rosier-Montus, M.F., Prades, C., Lemoine, C., Naudin, L., Denefle, P., Brewer, B., Duverger, N., Remaley, A. and Santamarina-Pojo, S.  
TITLE Regulatory nucleic acid sequences of the abcl gene  
JOURNAL Patent: WO 0183746-A 10 08-NOV-2001; Aventis Pharma S.A. (FR)

FEATURES  
source Location/Qualifiers  
1..9741  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 2650 a 2180 c 2290 g 2620 t 1 others  
ORIGIN

Query Match 41.6%; Score 92; DB 6; Length 9741;  
Best Local Similarity 100.0%; Pred. No. 1.7e-42;  
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 cttgtttttcccggttctgttttcccttcccttccggaagcttgcgaagggttagga 189  
|||||  
Db 1 CTTGTTTTTCCCGGTTCTGTTTCTCCCTTCTCCGAGAGGCTTGTCAGGGGTAGGA 60  
|||||

QY 190 gaaagagacgcaaacacaaaagtgaacaag 221  
|||||  
Db 61 GAAAGAGACGCAACACAAAAGTGGAACAG 92  
|||||

RESULT 30  
AX060715 1643 bp DNA linear PAT 22-JAN-2001  
LOCUS  
DEFINITION Sequence 3 from Patent WO0078972.  
ACCESSION AX060715  
VERSION AX060715.1 GI:12406104  
KEYWORDS

SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 1643)  
AUTHORS Lawn, R.M., Wade, D. and Garvin, M.  
TITLE Regulation with binding cassette transporter protein abcl  
JOURNAL Patent: WO 0078972-A 3 28-DEC-2000; CV THERAPEUTICS, INC. (US)

FEATURES  
source Location/Qualifiers  
1..1643  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 370 a 413 c 457 g 403 t  
ORIGIN

Query Match 41.2%; Score 91; DB 6; Length 1643;  
Best Local Similarity 100.0%; Pred. No. 6.4e-42;  
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gtaattgcagcagagtgagtgccggagcccgagagccgagagccgaccccttcttc 60  
|||||  
Db 1553 GTAATTGCGAGCGAGAGTGAGTGGCGCGGACCCGAGAGCCGACCCCTTCTCTC 1612  
|||||

QY 61 ccgggctgcggcagggcagggcgaggagctc 91  
|||||

Db 1613 CCGGGCTGCGGCGAGGCGAGGGGGGAGCTC 1643  
|||||

RESULT 31  
AX060894 1643 bp DNA linear PAT 22-JAN-2001  
LOCUS  
DEFINITION Sequence 3 from Patent WO0078971.  
ACCESSION AX060894  
VERSION AX060894.1 GI:12406271  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 1643)  
AUTHORS Lawn, R.M., Wade, D., Oram, J.F. and Garvin, M.  
TITLE Atp binding cassette transporter protein abcl polypeptides  
JOURNAL Patent: WO 0078971-A 3 28-DEC-2000; CV THERAPEUTICS, INC. (US)

FEATURES  
source Location/Qualifiers  
1..1643  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 370 a 413 c 457 g 403 t  
ORIGIN

Query Match 41.2%; Score 91; DB 6; Length 1643;  
Best Local Similarity 100.0%; Pred. No. 6.4e-42;  
Matches 91; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gtaattgcagcagagtgagtgccggagcccgagagccgagagccgaccccttcttc 60  
|||||  
Db 1553 GTAATTGCGAGCGAGAGTGAGTGGCGCGGACCCGAGAGCCGACCCCTTCTCTC 1612  
|||||

QY 61 ccgggctgcggcagggcagggcgaggagctc 91  
|||||

Db 1613 CCGGGCTGCGGCGAGGCGAGGGGGGAGCTC 1643  
|||||

RESULT 32  
AC021246/c 69570 bp DNA linear HTG 13-JUL-2000  
LOCUS  
DEFINITION Homo sapiens clone RP11-1N10, LOW-PASS SEQUENCE SAMPLING.  
ACCESSION AC021246  
VERSION AC021246.2 GI:9119882

KEYWORDS HTG: HTGS\_PHASE0.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 69570)  
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
TITLE Homo sapiens chromosome, clone RP11-1N10  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 69570)  
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
Anderson, S., Baldwin, J., Barna, N., Becker, R., Bieda, F.,  
Boguslavsky, L., Bouckgalter, B., Brown, A., Burkett, G., Castle, A.,  
Choepe, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,  
DeRellano, K., Dewar, K., Domino, M., Doyle, M., Fenestor, J.,  
Ferreira, P., FitzHugh, W., Forrest, C., Gage, D., Galagan, J.,  
Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,  
Landers, T., Lehoczy, J., Levine, R., Liu, C., Liu, G., Locke, K.,  
Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,  
McPheeters, R., Meldrum, J., Meneus, L., Morrow, J., Naylor, J.,  
Norman, C. H., O'Connor, T., O'Donnell, P., Olivari, T. M., Peterson, K.,  
Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,  
Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,  
Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J.,  
Zimmer, A. and Zody, W.  
TITLE Direct Submission  
JOURNAL Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
COMMENT On Jul 13, 2000 this sequence version replaced gl:6705871.  
All repeats were identified using RepeatMasker:  
Smit, A. F. A. & Green, P. (1996-1997).  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: L2512  
Center clone name: L\_N\_10  
-----  
\* NOTE: This record contains 73 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.  
\* 1 871: contig of 871 bp in length  
\* 972 971: gap of 100 bp  
\* 972 1834: contig of 863 bp in length  
\* 1835 1934: gap of 100 bp  
\* 1935 2804: contig of 870 bp in length  
\* 2805 2904: gap of 100 bp  
\* 2905 3745: contig of 841 bp in length  
\* 3746 3845: gap of 100 bp  
\* 3846 4696: contig of 851 bp in length  
\* 4697 4796: gap of 100 bp  
\* 4797 5640: contig of 844 bp in length  
\* 5641 5740: gap of 100 bp  
\* 5741 6540: contig of 800 bp in length  
\* 6541 6640: gap of 100 bp  
\* 6641 7509: contig of 869 bp in length  
\* 7510 7609: gap of 100 bp  
\* 7610 8479: contig of 870 bp in length  
\* 8480 8579: gap of 100 bp  
\* 8580 9430: contig of 851 bp in length  
\* 9431 9530: gap of 100 bp  
\* 9531 10376: contig of 846 bp in length  
\* 10377 10476: gap of 100 bp  
\* 10477 11322: contig of 846 bp in length  
\* 11323 11422: gap of 100 bp  
\* 11423 12302: contig of 880 bp in length  
\* 12303 12402: gap of 100 bp  
\* 12403 13280: contig of 878 bp in length  
\* 13281 13380: gap of 100 bp  
\* 13381 14241: contig of 861 bp in length  
\* 14242 14341: gap of 100 bp  
\* 14342 15196: contig of 855 bp in length  
\* 15197 15296: gap of 100 bp  
\* 15297 16123: contig of 827 bp in length  
\* 16124 16223: gap of 100 bp  
\* 16224 17072: contig of 849 bp in length  
\* 17073 17172: gap of 100 bp  
\* 17173 18041: contig of 869 bp in length  
\* 18042 18141: gap of 100 bp  
\* 18142 19009: contig of 868 bp in length  
\* 19010 19109: gap of 100 bp  
\* 19110 19966: contig of 857 bp in length  
\* 19967 20066: gap of 100 bp  
\* 20067 20921: contig of 855 bp in length  
\* 20922 21021: gap of 100 bp  
\* 21022 21865: contig of 844 bp in length  
\* 21866 21965: gap of 100 bp  
\* 21966 22832: contig of 867 bp in length  
\* 22833 22932: gap of 100 bp  
\* 22933 23780: contig of 848 bp in length  
\* 23781 23880: gap of 100 bp  
\* 23881 24733: contig of 853 bp in length  
\* 24734 24833: gap of 100 bp  
\* 24834 25670: contig of 837 bp in length  
\* 25671 25770: gap of 100 bp  
\* 25771 26621: contig of 851 bp in length  
\* 26622 26721: gap of 100 bp  
\* 26722 27576: contig of 855 bp in length  
\* 27577 27676: gap of 100 bp  
\* 27677 28532: contig of 856 bp in length  
\* 28533 28632: gap of 100 bp  
\* 28633 29492: contig of 860 bp in length  
\* 29493 29592: gap of 100 bp  
\* 29593 30455: contig of 863 bp in length  
\* 30456 30555: gap of 100 bp  
\* 30556 31410: contig of 855 bp in length  
\* 31411 31510: gap of 100 bp  
\* 31511 32368: contig of 858 bp in length  
\* 32369 32468: gap of 100 bp  
\* 32469 33312: contig of 844 bp in length  
\* 33313 33412: gap of 100 bp  
\* 33413 34268: contig of 856 bp in length  
\* 34269 34368: gap of 100 bp  
\* 34369 35204: contig of 836 bp in length  
\* 35205 35304: gap of 100 bp  
\* 35305 36156: contig of 852 bp in length  
\* 36157 36256: gap of 100 bp  
\* 36257 37128: contig of 872 bp in length  
\* 37129 37228: gap of 100 bp  
\* 37229 38083: contig of 855 bp in length  
\* 38084 38183: gap of 100 bp  
\* 38184 39031: contig of 848 bp in length  
\* 39032 39131: gap of 100 bp  
\* 39132 40006: contig of 875 bp in length  
\* 40007 40106: gap of 100 bp  
\* 40107 40967: contig of 861 bp in length  
\* 40968 41067: gap of 100 bp  
\* 41068 41913: contig of 846 bp in length  
\* 41914 42013: gap of 100 bp  
\* 42014 42824: contig of 811 bp in length  
\* 42825 42924: gap of 100 bp  
\* 42925 43776: contig of 852 bp in length  
\* 43777 43876: gap of 100 bp

\* 43877 44752: contig of 876 bp in length  
 \* 44753 44852: gap of 100 bp  
 \* 44853 45724: contig of 872 bp in length  
 \* 45725 45824: gap of 100 bp  
 \* 45825 46643: contig of 819 bp in length  
 \* 46644 46743: gap of 100 bp  
 \* 46744 47599: contig of 856 bp in length  
 \* 47600 47699: gap of 100 bp  
 \* 47700 48551: contig of 852 bp in length  
 \* 48552 48651: gap of 100 bp  
 \* 48652 49485: contig of 834 bp in length  
 \* 49486 49585: gap of 100 bp  
 \* 49586 50440: contig of 855 bp in length  
 \* 50441 50540: gap of 100 bp  
 \* 50541 51404: contig of 864 bp in length  
 \* 51405 51504: gap of 100 bp  
 \* 51505 52372: contig of 868 bp in length  
 \* 52373 52472: gap of 100 bp  
 \* 52473 53328: contig of 856 bp in length  
 \* 53329 53428: gap of 100 bp  
 \* 53429 54268: contig of 840 bp in length  
 \* 54269 54368: gap of 100 bp  
 \* 54369 55229: contig of 861 bp in length  
 \* 55230 55329: gap of 100 bp  
 \* 55330 56197: contig of 868 bp in length  
 \* 56198 56297: gap of 100 bp  
 \* 56298 57163: contig of 866 bp in length  
 \* 57164 57263: gap of 100 bp  
 \* 57264 58130: contig of 867 bp in length  
 \* 58131 58230: gap of 100 bp  
 \* 58231 59082: contig of 852 bp in length  
 \* 59083 59182: gap of 100 bp  
 \* 59183 60020: contig of 838 bp in length  
 \* 60021 60120: gap of 100 bp  
 \* 60121 60983: contig of 863 bp in length  
 \* 60984 61083: gap of 100 bp  
 \* 61084 61935: contig of 852 bp in length  
 \* 61936 62035: gap of 100 bp  
 \* 62036 62866: contig of 831 bp in length  
 \* 62867 62966: gap of 100 bp  
 \* 62967 63827: contig of 861 bp in length  
 \* 63828 63927: gap of 100 bp  
 \* 63928 64783: contig of 856 bp in length  
 \* 64784 64883: gap of 100 bp  
 \* 64884 65740: contig of 857 bp in length  
 \* 65741 65840: gap of 100 bp  
 \* 65841 66684: contig of 844 bp in length  
 \* 66685 66784: gap of 100 bp  
 \* 66785 67651: contig of 867 bp in length  
 \* 67652 67751: gap of 100 bp

Query Match 21.3%; Score 47; DB 2; Length 69570;  
 Best Local Similarity 100.0%; Pred. No. 3.7e-16;  
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 caggcgaggagctccgcgcaccacagagccggttctcagagcgct 123  
 |||||  
 Db 67335 CAGGCGGGAGGTCCCGCCACACAGAGCCGTTCTCAGGCGGCT 67289

RESULT 33  
 AC097517/c  
 LOCUS Homo sapiens chromosome 2 clone RP11-485017, complete sequence.  
 DEFINITION AC097517  
 ACCESSION AC097517  
 VERSION AC097517.3 GI:17998642  
 KEYWORDS HTG.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 92227)  
 AUTHORS Waterston, R.H.

TITLE The sequence of Homo sapiens clone  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 92227)  
 AUTHORS Waterston, R.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (18-OCT-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
 REFERENCE 3 (bases 1 to 92227)  
 AUTHORS Waterston, R.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (29-DEC-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
 REFERENCE 4 (bases 1 to 92227)  
 AUTHORS Waterston, R.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (03-JAN-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
 COMMENT On Dec 29, 2001 this sequence version replaced gi:16902004.

----- Genome Center -----  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: http://genome.wustl.edu/gsc/index.shtml  
 Contact: submissions@watson.wustl.edu  
 ----- Project Information -----  
 Center project name: H\_NH0485017  
 Drafting center: WIBR  
 -----

FEATURES  
 source

Location/Qualifiers  
 1..92227  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="2"  
 /clone="RP11-485017"  
 BASE COUNT 27775 a 18464 c 18146 g 27842 t  
 ORIGIN

Query Match 9.0%; Score 20; DB 9; Length 92227;  
 Best Local Similarity 100.0%; Pred. No. 2.4;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 149 tggttttcccttctccgg 168  
 |||||  
 Db 73615 TGTTCCTCCCTCTCCGG 73596

RESULT 34  
 AC022563

LOCUS Homo sapiens clone RP11-3L23, LOW-PASS SEQUENCE SAMPLING.  
 DEFINITION AC022563  
 ACCESSION AC022563  
 VERSION AC022563.2 GI:9121084  
 KEYWORDS HTG; HTGS\_PHASE0.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 149710)  
 AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
 TITLE Unpublished  
 JOURNAL Homo sapiens chromosome, clone RP11-3L23  
 REFERENCE 2 (bases 1 to 149710)  
 AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Beda, F., Bozulavkiy, I., Boukhgalter, B., Brown, A., Burkett, G., Castle, A., Choepel, I., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearrellano, K., Dewar, K., Domino, M., Doyle, M., Fenestor, J., Ferrelli, P., FitzHugh, W., Forrest, C., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,



Howland, J.C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,  
Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G., Locke, K.,  
Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,  
McPheeters, R., Meldrum, J., Meneus, L., Morrow, J., Naylor, J.,  
Norman, C.H., O'Connor, T., O'Donnell, P., Olivari, T.M., Peterson, K.,  
Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,  
Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,  
Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W.J.,  
Zimmer, A., and Zody, M.

## TITLE

## JOURNAL

## COMMENT

Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jul 13, 2000 this sequence version replaced gi:6910806.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project information

Center project name: L2771

Center clone name: 3\_L\_23

\* NOTE: This record contains 171 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

1 682: contig of 682 bp in length  
683 782: gap of 100 bp  
783 1492: contig of 710 bp in length  
1493 1592: gap of 100 bp  
1593 2305: contig of 713 bp in length  
2306 2405: gap of 100 bp  
2406 3113: contig of 708 bp in length  
3114 3213: gap of 100 bp  
3214 3867: contig of 654 bp in length  
3868 3967: gap of 100 bp  
3968 4646: contig of 679 bp in length  
4647 4746: gap of 100 bp  
4747 5416: contig of 670 bp in length  
5417 5516: gap of 100 bp  
5517 6193: contig of 677 bp in length  
6194 6293: gap of 100 bp  
6294 6968: contig of 675 bp in length  
6969 7068: gap of 100 bp  
7069 7757: contig of 689 bp in length  
7758 7857: gap of 100 bp  
7858 8557: contig of 700 bp in length  
8558 8657: gap of 100 bp  
8658 9336: contig of 679 bp in length  
9337 9436: gap of 100 bp  
9437 10124: contig of 688 bp in length  
10125 10224: gap of 100 bp  
10225 10879: contig of 655 bp in length  
10880 10979: gap of 100 bp  
10980 11663: contig of 684 bp in length  
11664 11763: gap of 100 bp  
11764 12433: contig of 670 bp in length  
12434 12533: gap of 100 bp  
12534 13209: contig of 676 bp in length  
13210 13309: gap of 100 bp  
13310 13989: contig of 680 bp in length  
13990 14089: gap of 100 bp

14090 14798: contig of 709 bp in length  
14799 14898: gap of 100 bp  
14899 15571: contig of 673 bp in length  
15572 15671: gap of 100 bp  
15672 16387: contig of 716 bp in length  
16388 16487: gap of 100 bp  
16488 17198: contig of 711 bp in length  
17199 17298: gap of 100 bp  
17299 18008: contig of 710 bp in length  
18009 18108: gap of 100 bp  
18109 18772: contig of 664 bp in length  
18773 18872: gap of 100 bp  
18873 19555: contig of 683 bp in length  
19556 19655: gap of 100 bp  
19656 20320: contig of 665 bp in length  
20321 20420: gap of 100 bp  
20421 21112: contig of 692 bp in length  
21113 21212: gap of 100 bp  
21213 21885: contig of 673 bp in length  
21886 21985: gap of 100 bp  
21986 22671: contig of 686 bp in length  
22672 22771: gap of 100 bp  
22772 23451: contig of 680 bp in length  
23452 23551: gap of 100 bp  
23552 24267: contig of 716 bp in length  
24268 24367: gap of 100 bp  
24368 25073: contig of 706 bp in length  
25074 25173: gap of 100 bp  
25174 25867: contig of 694 bp in length  
25868 25967: gap of 100 bp  
25968 26669: contig of 702 bp in length  
26670 26769: gap of 100 bp  
26770 27445: contig of 676 bp in length  
27446 27545: gap of 100 bp  
27546 28224: contig of 679 bp in length  
28225 28324: gap of 100 bp  
28325 29002: contig of 678 bp in length  
29003 29102: gap of 100 bp  
29103 29785: contig of 683 bp in length  
29786 29885: gap of 100 bp  
29886 30571: contig of 686 bp in length  
30572 30671: gap of 100 bp  
30672 31357: contig of 686 bp in length  
31358 31457: gap of 100 bp  
31458 32168: contig of 711 bp in length  
32169 32268: gap of 100 bp  
32269 32951: contig of 683 bp in length  
32952 33051: gap of 100 bp  
33052 33756: contig of 705 bp in length  
33757 33856: gap of 100 bp  
33857 34566: contig of 710 bp in length  
34567 34666: gap of 100 bp  
34667 35376: contig of 710 bp in length  
35377 35476: gap of 100 bp  
35477 36143: contig of 667 bp in length  
36144 36243: gap of 100 bp  
36244 36924: contig of 681 bp in length  
36925 37024: gap of 100 bp  
37025 37701: contig of 677 bp in length  
37702 37801: gap of 100 bp  
37802 38493: contig of 692 bp in length  
38494 38593: gap of 100 bp  
38594 39274: contig of 681 bp in length  
39275 39374: gap of 100 bp  
39375 40071: contig of 697 bp in length  
40072 40171: gap of 100 bp  
40172 40870: contig of 699 bp in length  
40871 40970: gap of 100 bp  
40971 41659: contig of 689 bp in length  
41660 41759: gap of 100 bp  
41760 42434: contig of 675 bp in length  
42435 42534: gap of 100 bp  
42535 43207: contig of 673 bp in length

\* 43208 43307: gap of 100 bp  
 \* 43308 43980: contig of 673 bp in length  
 \* 43981 44080: gap of 100 bp  
 \* 44081 44768: contig of 688 bp in length  
 \* 44769 44868: gap of 100 bp  
 \* 44869 45556: contig of 688 bp in length  
 \* 45557 45656: gap of 100 bp  
 \* 45657 46348: contig of 692 bp in length  
 \* 46349 46448: gap of 100 bp  
 \* 46449 47143: contig of 695 bp in length  
 \* 47144 47243: gap of 100 bp  
 \* 47244 47943: contig of 700 bp in length  
 \* 47944 48043: gap of 100 bp  
 \* 48044 48727: contig of 684 bp in length  
 \* 48728 48837: gap of 100 bp  
 \* 48838 49540: contig of 713 bp in length  
 \* 49541 49640: gap of 100 bp  
 \* 49641 50332: contig of 692 bp in length  
 \* 50333 50432: gap of 100 bp  
 \* 50433 51127: contig of 695 bp in length  
 \* 51128 51227: gap of 100 bp  
 \* 51228 51915: contig of 688 bp in length  
 \* 51916 52015: gap of 100 bp  
 \* 52016 52678: contig of 663 bp in length  
 \* 52679 52778: gap of 100 bp  
 \* 52779 53460: contig of 682 bp in length  
 \* 53461 53560: gap of 100 bp  
 \* 53561 54251: contig of 691 bp in length  
 \* 54252 54351: gap of 100 bp  
 \* 54352 54921: contig of 570 bp in length  
 \* 54922 55021: gap of 100 bp  
 \* 55022 55702: contig of 681 bp in length  
 \* 55703 55802: gap of 100 bp

Query Match 9.0%; Score 20; DB 2; Length 149710;  
 Best Local Similarity 100.0%; Pred. No. 2.4;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 170 agcgtgtcaagggttagga 189  
 |||||

Db 55886 AGGCTTCAAGGGTAGGA 55905

# RESULT 35

AC027453  
 LOCUS Homo sapiens chromosome 10 clone RP11-524H12 map 10, WORKING DRAFT  
 DEFINITION AC027453.3 GI:8077127  
 ACCESSION AC027453  
 VERSION HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
 KEYWORDS human.  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 162496)  
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N.,  
 Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,  
 Boguski,M., Bouckgeer,B., Brown,A., Burkett,G.,  
 Campopiano,A., Cooke,A., Choquel,Y., Colangelo,M., Collins,S.,  
 Collymore,A., Cooke,P., DeAfrillano,K., Dewar,K., Diaz,J.S.,  
 Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,  
 Galagan,J., Gardyna,S., Ginde,S., Govette,M., Graham,L.,  
 Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,  
 Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,  
 Kiehl,J., LaRoque,K., Lamazares,R., Landers,T., Lehotzky,J.,  
 Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,  
 McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,  
 Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,  
 Nelson,S., Nuzum,A., O'Connor,T., O'Donnell,P.,  
 O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,  
 Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,  
 Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,  
 Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
 Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,  
 Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,  
 Young,G., Zainoun,J., Zimmer,A. and Zody,M.

Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
 O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,  
 Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,  
 Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,  
 Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
 Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,  
 Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,  
 Young,G., Zainoun,J., Zimmer,A. and Zody,M.

## TITLE JOURNAL COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: <http://www-seq.wi.mit.edu>  
 Contact: [sequence.submissions@genome.wi.mit.edu](mailto:sequence.submissions@genome.wi.mit.edu)

----- Project Information  
 Center project name: L8240  
 Center clone name: 524\_H12

----- Summary Statistics  
 Sequencing vector: M13; M77815; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.960731  
 Consensus quality: 154128 bases at least Q40  
 Consensus quality: 158489 bases at least Q30  
 Consensus quality: 160079 bases at least Q20  
 Insert size: 163000; agarose-fp  
 Insert size: 160996; sum-of-contents  
 Quality coverage: 4.7 in Q20 bases; agarose-fp  
 Quality coverage: 4.8 in Q20 bases; sum-of-contents

----- NOTE: This is a 'working draft' sequence. It currently  
 consists of 16 contigs. The true order of the pieces  
 is not known and their order in this sequence record is  
 arbitrary. Gaps between the contigs are represented as  
 runs of N, but the exact sizes of the gaps are unknown.  
 This record will be updated with the finished sequence  
 as soon as it is available and the accession number will  
 be preserved.

1  
 882 981: contig of 881 bp in length  
 982 4366: gap of 100 bp  
 4367 4466: gap of 100 bp  
 4467 7420: contig of 2954 bp in length  
 7421 7520: gap of 100 bp  
 7521 11683: contig of 4163 bp in length  
 11684 11783: gap of 100 bp  
 11784 15264: contig of 3481 bp in length  
 15265 15364: gap of 100 bp  
 15365 20920: contig of 5556 bp in length  
 20921 21020: gap of 100 bp  
 21021 27969: contig of 6949 bp in length  
 27970 28069: gap of 100 bp  
 28070 36585: contig of 8516 bp in length  
 36586 36685: gap of 100 bp  
 36686 47838: contig of 11153 bp in length  
 47839 47938: gap of 100 bp  
 47939 57920: contig of 9982 bp in length  
 57921 58020: gap of 100 bp  
 58021 70481: contig of 12461 bp in length  
 70482 70581: gap of 100 bp  
 70582 80819: contig of 10238 bp in length  
 80820 80919: gap of 100 bp  
 80920 91740: contig of 10821 bp in length  
 91741 91840: gap of 100 bp  
 91841 113943: contig of 22103 bp in length  
 113944 114043: gap of 100 bp  
 114044 137119: contig of 23076 bp in length  
 137120 137219: gap of 100 bp

\* 137220 162496: contig of 25277 bp in length.

## FEATURES

Location/Qualifiers  
 source 1. .162496  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="10"  
 /map="10"  
 /clone\_lib="RPCI-11 Human Male BAC"  
 misc\_feature 1. .881  
 /note="assembly\_fragment  
 clone\_end:T7  
 vector\_side:left"  
 misc\_feature 982. .4366  
 /note="assembly\_fragment"  
 misc\_feature 4467. .7420  
 /note="assembly\_fragment"  
 misc\_feature 7521. .11683  
 /note="assembly\_fragment"  
 misc\_feature 11784. .15264  
 /note="assembly\_fragment  
 clone\_end:SP6  
 vector\_side:right"  
 misc\_feature 15365. .20920  
 /note="assembly\_fragment"  
 misc\_feature 21021. .27969  
 /note="assembly\_fragment"  
 misc\_feature 28070. .36585  
 /note="assembly\_fragment"  
 misc\_feature 36686. .47838  
 /note="assembly\_fragment"  
 misc\_feature 47939. .57920  
 /note="assembly\_fragment"  
 misc\_feature 58021. .70481  
 /note="assembly\_fragment"  
 misc\_feature 70582. .80819  
 /note="assembly\_fragment"  
 misc\_feature 80920. .91740  
 /note="assembly\_fragment"  
 misc\_feature 91841. .113943  
 /note="assembly\_fragment"  
 misc\_feature 114044. .137119  
 /note="assembly\_fragment"  
 misc\_feature 137220. .162496  
 /note="assembly\_fragment"  
 BASE COUNT 47419 a 33487 c 33493 g 46593 t 1504 others  
 ORIGIN

Query Match 9.0%; Score 20; DB 2; Length 162496;

Best Local Similarity 100.0%; Pred. No. 2.4; Indels 0; Gaps 0;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 149 tgttttcccttcttcg 168

Db 113242 TGTTCCTCCCTCTCCGG 113261

## RESULT 36

## AC012325

LOCUS 201938 bp DNA linear HTG 25-APR-2001  
 DEFINITION Homo sapiens chromosome 16 clone RP11-93H5, WORKING DRAFT SEQUENCE,  
 7 ordered pieces.

## ACCESSION AC012325

VERSION AC012325.6 GI:9954638

KEYWORDS HTG; HTGS\_PHASE2; HTGS\_DRAFT; HTGS\_ACTIVEFIN.

## SOURCE human.

## ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 201938)

AUTHORS DOE Joint Genome Institute.

TITLE Sequencing of Human Chromosome 16

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## Unpublished

2 (bases 1 to 201938)

DOE Joint Genome Institute.

Direct Submission

Submitted (23-OCT-1999) Production Sequencing Facility, DOE Joint  
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

On Aug 31, 2000 this sequence version replaced gi:7690207.

Sequence Quality Assessment:

This entry has been annotated with sequence quality

estimates computed by the Phrap assembly program.

All manually edited bases have been reduced to quality zero.

Quality levels above 40 are expected to have less than

1 error in 10,000 bp.

Base-by-base quality values are not generally visible from the

GenBank flat file format but are available as part

of this entry's ASN.1 file.

-----

Sequence Quality Assessment:

This entry has been annotated with sequence quality

estimates computed by the Phrap assembly program.

All manually edited bases have been reduced to quality zero.

Quality levels above 40 are expected to have less than

1 error in 10,000 bp.

Base-by-base quality values are not generally visible from the

GenBank flat file format but are available as part

of this entry's ASN.1 file.

-----

\* NOTE: This is a 'working draft' sequence. It currently

consists of 7 contigs. Gaps between the contigs

are represented as runs of N. The order of the pieces

is believed to be correct as given, however the sizes

of the gaps between them are based on estimates that have

provided by the submitter.

\* This sequence will be replaced

\* by the finished sequence as soon as it is available and

\* the accession number will be preserved.

\* 1 33343: contig of 33343 bp in length

\* 33344 33443: gap of unknown length

\* 33444 40628: contig of 7185 bp in length

\* 40629 40728: gap of unknown length

\* 40729 143349: contig of 102621 bp in length

\* 143350 143449: gap of unknown length

\* 143450 144499: contig of 1050 bp in length

\* 144500 144599: gap of unknown length

\* 144600 162264: contig of 17665 bp in length

\* 162265 162364: gap of unknown length

\* 162365 187856: contig of 25492 bp in length

\* 187857 187956: gap of unknown length

\* 187957 201938: contig of 13982 bp in length.

Location/Qualifiers

1. .201938

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="16"

/clone="RP11-93H5"

BASE COUNT 54511 a 50297 c 47816 g 48714 t 600 others

## ORIGIN

## Query Match

Best Local Similarity 9.0%; Score 20; DB 2; Length 201938;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 63 gggctgcgcaggcaggc 82

Db 186942 GGGCTGCGCAGGCAGGC 186961

## RESULT 37

## AC097253/c

## LOCUS

## DEFINITION

## SEQUENCE, 12 unordered pieces.

228703 bp DNA linear HTG 20-DEC-2001

Rattus norvegicus chromosome SA clone CH230-152G15, WORKING DRAFT

AC097253 AC097253.3 GI:17973859  
 HTG: HTGS\_PHASE1; HTGS\_DRAFT.  
 SOURCE Norway rat.  
 ORGANISM Rattus norvegicus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

AUTHORS 1 (bases 1 to 228703)  
 Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C., Alsbrooks, S.L., Amarantunga, H.C., Are, J.R., Banks, T., Barbara, J., Benton, J., Blimege, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Channon, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Haviak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hognes, M., Hollaway, C., Hollins, B., Honsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseg, H., Lozado, R.J., Lu, X., Lucier, R., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhinney, E., McLeod, M.P., Meador, M., Mel, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenko, S., Oghu, M., Okwona, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalobos, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G. and Gibbs, R.

Direct Submission  
 Unpublished  
 2 (bases 1 to 228703)  
 Worley, K.C.

TITLE JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

COMMENT Submitted (13-OCT-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
 On Dec 20, 2001 this sequence version replaced gi:17062530.  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
 ----- Project Information  
 Center project name: GIOR  
 Center clone name: CH230-152G15  
 ----- Summary Statistics  
 Assembly program: Phrap; version 0.990329First call to findPhraplist  
 Consensus quality: 221817 bases at least Q40  
 Consensus quality: 223045 bases at least Q30  
 Consensus quality: 223940 bases at least Q20  
 Estimated insert size: 224711; sum-of-contigs estimation  
 Quality coverage: 0x in Q20 bases; agarose-ip estimation

Quality coverage: 6.4x in Q20 bases; sum-of-contigs estimation  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 12 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1 40920: contig of 40920 bp in length  
 \* 40921 41020: gap of unknown length  
 \* 41021 73829: contig of 32809 bp in length  
 \* 73830 73929: gap of unknown length  
 \* 73930 100142: contig of 26213 bp in length  
 \* 100143 100242: gap of unknown length  
 \* 100243 126981: contig of 26739 bp in length  
 \* 126982 127081: gap of unknown length  
 \* 127082 150119: contig of 23038 bp in length  
 \* 150120 150219: gap of unknown length  
 \* 150220 172428: contig of 22209 bp in length  
 \* 172429 172528: gap of unknown length  
 \* 172529 191512: contig of 18983 bp in length  
 \* 191513 209925: contig of 18314 bp in length  
 \* 209926 210025: gap of unknown length  
 \* 210026 219870: contig of 9845 bp in length  
 \* 219871 225401: contig of 5431 bp in length  
 \* 225402 225501: gap of unknown length  
 \* 225502 227014: contig of 1513 bp in length  
 \* 227015 227114: gap of unknown length  
 \* 227115 228703: contig of 1589 bp in length.  
 \* Location/Qualifiers  
 1..228703  
 /organism="Rattus norvegicus"  
 /db\_xref="taxon:10116"  
 /chromosome="SA"  
 /clone="CH230-152G15"  
 BASE COUNT 65603 a 46098 c 46478 g 69392 t 1132 others  
 ORIGIN

Query Match 9.0%; Score 20; DB 2; Length 228703;  
 Best Local Similarity 100.0%; Pred. No. 2.4;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 ttctgtttttctccctctc 165  
 |||||  
 DB 82733 TTCTGTTTTCCTCCCTCTC 82714

RESULT 38  
 A39868/c  
 LOCUS  
 DEFINITION Sequence 2 from Patent EP0614981.  
 ACCESSION A39868  
 VERSION A39868.1 GI:2296098  
 KEYWORDS  
 SOURCE unidentified.  
 ORGANISM unidentified.  
 unclassified.  
 REFERENCE 1 (bases 1 to 735)  
 AUTHORS Dembic, Z., Garotta, G. and Gentz, R. H.  
 TITLE Chimeric human interferon-gamma-receptor/immunoglobulin polypeptides  
 JOURNAL Hoffmann LA ROCHE (CH)  
 COMMENT Patent: EP 0614981-A 2 14-SEP-1994;  
 Other publication JP 6319552 941122  
 Other publication NZ 250997 951026  
 Other publication CA 2114168 940906

Other publication CN 1094092 941026  
Other publication AU 5647894 940908  
Other publication ZA 9401333 940906.

## FEATURES

Location/Qualifiers  
1. 735

/organism="unidentified"

/db\_xref="taxon:32644"

223 a 135 c 170 g 207 t

BASE COUNT  
ORIGIN

Query Match 8.6%; Score 19; DB 6; Length 735;  
Best Local Similarity 100.0%; Pred. No. 8.7;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 178 caaggggtaggagaagag 196

|||||

Db 28 CAAGGGGTAGGAGAAAG 10

## RESULT 39

AF056979/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

Location/Qualifiers

1. 1641

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="YANI"

/sex="female"

/cell\_line="HVS transformed"

/cell\_type="lymphocyte"

/tissue\_type="blood"

22. 1491

/codon\_start=1

/product="interferon-gamma receptor"

/protein\_id="AAK30623.1"

/db\_xref="GI:13562049"

/translation="MALLFLPLVMQGVSRAMGTADLGPSSVPTPTNTVTIESYNMNP

IVWYQIMQVPVFTVEGKNGVKNSEWIDACINISHHYCNISDHVGDPSNSLWVR

KALNQKESAKSEFAVCVRGKIGPKLDIRKEKOIMIDIFPHSVFVNGDQEV

YDPETTCYIRVNYVYVRNGSEIQYKILTKEDCCDELOCOLAIPVSLNSQYCSAE

GLVHWVPTKSRKVCITIFNSSKLSGIMIPVVAALLFLVLSLVCIFCKINPL

REKSIILPKSLISVRSATLETKPSKYSLITVQPSFLEKVCCEPLSPATVPM

ALNSVHRSNCSHRSNGFTDSSCLSHSLSDSEPPNNKGEIKTEGQELITVIK

APTSPGYDKPHVLDLLVDDSGKSLIGYRTEDSKFS"

508 a 320 c 352 g 461 t

BASE COUNT

ORIGIN

Location/Qualifiers

1. 2064

source

Location/Qualifiers

1. 2064

source

Query Match 8.6%; Score 19; DB 9; Length 1641;  
Best Local Similarity 100.0%; Pred. No. 8.8;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 178 caaggggtaggagaagag 196

|||||

Db 49 CAAGGGGTAGGAGAAAG 31

## RESULT 40

SIVAGMGAGD

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

FEATURES

source

Location/Qualifiers

1. 1773

/organism="Simian immunodeficiency virus"

/proviral

/db\_xref="taxon:11723"

/clone="agm49gag"

/haplotype="na"

/cell\_line="CEMss"

195. 1773

/genes="gag"

195. 1773

/genes="gag"

594 a 375 c 480 g 324 t

BASE COUNT

ORIGIN

Query Match 8.6%; Score 19; DB 14; Length 1773;  
Best Local Similarity 100.0%; Pred. No. 8.8;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 172 gcttgcagggtaggag 190

|||||

Db 1265 GCTTGTCAAGGGTAGGAG 1283

## RESULT 41

HUMIFNRG/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

FEATURES

source

Location/Qualifiers

1. 2064

source

Location/Qualifiers

1. 2064

source

HUMIFNRG 2064 bp mRNA linear PRI 08-NOV-1994  
Human interferon-gamma receptor mRNA, complete cds.

J03143

J03143.1 GI:184650

interferon receptor.

Human lymphoid tissue cell line Raji, cDNA to mRNA.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 2064)

Agnet,M., Dembic,Z. and Merlin,G.

Molecular cloning and expression of the human interferon-gamma

receptor

Cell 55 (2), 273-280 (1988)

89003065

Draft entry and computer-readable sequence for [1] kindly provided

by M.Aguet, 08-SEP-1988.

Location/Qualifiers

1. 2064

source

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="6q23-q24"  
<1..2064  
/gene="IFNGR1"  
/product="IFNR-gamma mRNA"  
1..2064  
/gene="IFNGR1"  
49..1518  
/note="interferon-gamma receptor"  
/codon\_start=1  
/protein\_id="AA52731.1"  
/db\_xref="GI:306915"  
/db\_xref="GDB:G00-120-688"  
/translation="MALLFLPLVMQGVSRAMGTADLGPSSVPTNTVIESYNNMP  
IVWEYQIMQPVFTVEVKNYGVKNSWIDACINSHHYCNISDRHVGDPNSLWVRV  
KARVGKESAYAKSEFAVCRDGIKPPKLDIRKEKQIMIDIFHPSVFNQDEQVD  
YDPETCYIRVYVYRMNGSEIQKILTKQEDDCDEIQCOLAIPVSSNQYCVSAE  
GVLHWGVTTEKKEVCITIFNSIKGSLWIPVAAALLFLVLSLFCFYIKKINPL  
KESIIILPKSLISVRSATLETPEKSYSLITSYQFSLKEVCEELSPATVPGM  
HTEDNPKVTEELSSITEVVTTEENIPDVVPGSHLPIERESSPLSSNQSEPGSI  
ALNSYHSRNCSESDSRNGFDTSSCLSHSLSDSEFPNPKNGEIKTEQELITIVK  
APTSFGYDKPHVLVDLLVDDSGKESLIGYRPTEDSKEFS"  
BASE COUNT 639 a 383 c 426 g 616 t  
ORIGIN 1 bp upstream of EcoRI site; chromosome 6q15-q21.

Query Match 8.6%; Score 19; DB 9; Length 2064;  
Best Local Similarity 100.0%; Pred. No. 8.8;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 178 caaggggtaggagaaag 196  
|||||  
Db 76 CAAGGGGTAGGAGAAAG 58

RESULT 42  
BC005333/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REMARK  
COMMENT

BC005333  
Homo sapiens, interferon gamma receptor 1, clone MGC:12420  
IMAGE:3950528, mRNA, complete cds.  
BC005333  
BC005333.1 GI:13529118  
MGC.  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 2127)  
Strausberg, R.  
Direct Submission  
Submitted (27-MAR-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgapsb-r@mail.nih.gov](mailto:cgapsb-r@mail.nih.gov)  
Tissue Procurement: CLONTECH  
CDNA Library Preparation: CLONTECH Laboratories, Inc.  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www-shgc.stanford.edu>  
Contact: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
R. M.  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 16 Row: k Column: 12  
This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 5419821.  
FEATURES  
source  
Location/Qualifiers  
1..2127  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/db\_xref="LocusID:3459"  
/clone="MGC:12420 IMAGE:3950528"  
/tissue\_type="prostate"  
/clone\_lib="NIH\_MGC\_83"  
/lab\_host="DH10B"  
/note="Vector: pDNR-LIB"  
83..1552  
/codon\_start=1  
/product="interferon gamma receptor 1"  
/protein\_id="AAH05333.1"  
/db\_xref="GI:13529119"  
/translation="MALLFLPLVMQGVSRAMGTADLGPSSVPTNTVIESYNNMP  
IVWEYQIMQPVFTVEVKNYGVKNSWIDACINSHHYCNISDRHVGDPNSLWVRV  
KARVGKESAYAKSEFAVCRDGIKPPKLDIRKEKQIMIDIFHPSVFNQDEQVD  
YDPETCYIRVYVYRMNGSEIQKILTKQEDDCDEIQCOLAIPVSSNQYCVSAE  
GVLHWGVTTEKKEVCITIFNSIKGSLWIPVAAALLFLVLSLFCFYIKKINPL  
KESIIILPKSLISVRSATLETPEKSYSLITSYQFSLKEVCEELSPATVPGM  
HTEDNPKVTEELSSITEVVTTEENIPDVVPGSHLPIERESSPLSSNQSEPGSI  
ALNSYHSRNCSESDSRNGFDTSSCLSHSLSDSEFPNPKNGEIKTEQELITIVK  
APTSFGYDKPHVLVDLLVDDSGKESLIGYRPTEDSKEFS"  
BASE COUNT 671 a 391 c 445 g 620 t  
ORIGIN

Query Match 8.6%; Score 19; DB 9; Length 2127;  
Best Local Similarity 100.0%; Pred. No. 8.8;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 178 caaggggtaggagaaag 196  
|||||  
Db 110 CAAGGGGTAGGAGAAAG 92

RESULT 43  
A30438/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
BASE COUNT  
ORIGIN

A30438  
H.sapiens pBABE Sact/Asp181 fragment.  
A30438  
A30438.1 GI:1567031  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 2184)  
Fountoulakis, M., Garotta, G. and Stueber, D.  
Soluble interferon-gamma receptors and methods for their production  
Patent: EP 0393502-A 1 24-OCT-1990;  
F. HOFFMANN-LA ROCHE AG  
Location/Qualifiers  
1..2184  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 688 a 413 c 452 g 631 t  
ORIGIN

Query Match 8.6%; Score 19; DB 6; Length 2184;  
Best Local Similarity 100.0%; Pred. No. 8.8;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 178 caaggggtaggagaaag 196  
|||||  
Db 112 CAAGGGGTAGGAGAAAG 94

RESULT 44  
STHK022